Oppression and the Limits of Individual Moral Progress

Cameron Charles Evans

*Washington University in St. Louis*

Follow this and additional works at: [https://openscholarship.wustl.edu/art_sci_etds](https://openscholarship.wustl.edu/art_sci_etds)

Part of the Philosophy Commons

**Recommended Citation**

[https://openscholarship.wustl.edu/art_sci_etds/2412](https://openscholarship.wustl.edu/art_sci_etds/2412)

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
# Table of Contents

**Acknowledgments**  

**Abstract**  

**Preface**  

**Chapter 1 - Can We Fix Our Psychological Biases?**  
1.1 - Introduction  
1.2 - The Problem with Implicit Biases  
  1.2.1 - Automatic, Unreflective, and Small  
  1.2.2 - Avoiding Triggers  
  1.2.3 - Counter-Conditioning  
1.3 - The Problem with Other Kinds of Biases  
  1.3.1 - Empathy Gap  
  1.3.2 - Identity and Group Conflict  
1.4 - The Problem of Social and Situational Influence  
1.5 - Conclusion  

**Chapter 2 - Can Social Practices Render Us Morally Ignorant?**  
2.1 - Introduction  
2.2 - Structural Ignorance Can Excuse Wrongdoing  
  2.2.1 - Responsibility for Wrongdoing  
  2.2.2 - The “Brainwashing” Excuse  
  2.2.3 - Responsibility for Ignorance  
2.3 - Objections  
  2.3.1 - But, Isn’t Moral Knowledge Easy?  
  2.3.2 - Aren’t They Just Willfully Ignorant?  
  2.3.3 - But Aren’t They Blameworthy Anyway?  
2.4 - Conclusion: Beyond Brainwashing  

**Chapter 3 - Do Our Social Identities Limit What We Can Know?**  
3.1 - Introduction  
3.2 - Biological Advantage  
3.3 - Social Advantage  
  3.3.1 - The Epistemic Value of Suffering  
  3.3.2 - Regularities in Experience  
3.4 - Lack of Disadvantage  
  3.4.1 - Masculine Biases and Prejudices
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2 - Vicious Men</td>
<td>82</td>
</tr>
<tr>
<td>3.4.3 - Vicious Societies</td>
<td>87</td>
</tr>
<tr>
<td>3.5 - Conclusion</td>
<td>90</td>
</tr>
<tr>
<td><strong>Conclusions: The Limits of Individual Moral Progress</strong></td>
<td>92</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td>100</td>
</tr>
</tbody>
</table>
Acknowledgments

I would like to give special thanks to Ron Mallon for his unwavering support in the completion of my dissertation and to Allan Hazlett for providing substantive feedback on early drafts. I am grateful for the patience of each of my committee members while I completed my work during the COVID pandemic and while starting a new job. I would also like to thank the Chancellor’s Graduate Fellowship for supporting me socially and financially from the start of my graduate career.

Cameron C. Evans

Washington University in St. Louis

May 2021
ABSTRACT OF THE DISSERTATION

Oppression and the Limits of Individual Moral Progress

by

Cameron Charles Evans

Doctor of Philosophy in Philosophy-Neuroscience-Psychology

Washington University in St. Louis, 2021

Ron Mallon, Chair

Systemic social problems are the product of complex interactions among institutions, policies, biases, stereotypes, and false beliefs. While we may be pessimistic about our abilities to combat social change, we may remain optimistic about our abilities as individuals to achieve moral growth. However, in societies organized by socially unjust practices, our individual capacities for improvement may be stunted. My primary concern is to identify key obstacles to our moral development.

I argue that "self-help" methods for reducing our social biases and prejudices are ineffective. Our biases interact in complex ways with each other, and with the environment, in ways that complicate our individual efforts for improvement. Moreover, I argue our abilities to identify and correct our false beliefs can be hindered by learned social practices. I argue such an ignorance can explain differences in epistemic success across social identities. For example, sexist social practices might encourage men to be willfully ignorant, or to develop epistemic vice.
Nonetheless, men making genuine attempts to understand sexism may characteristically fail to correct their ignorance due to pervasive sexist practices for conceptualizing social life and sexist epistemic practices.

If we fail to understand the factors that get in the way of our personal moral development, we may grow complacent with the illusion of moral growth and we may demonize others who need a helping hand. With a better understanding of the limits on our moral growth, we can start to think more productively about what we can actually do to affect positive moral change in ourselves and in the world.
Preface

In societies organized by socially unjust practices, our individual efforts to be better ourselves may be systematically thwarted. This claim may fail to surprise anyone who has attempted to fight against oppression.

Advocates for social change face a Gordian knot of institutions, policies, biases, stereotypes, and false beliefs that interact in complex ways and reinforce the status quo. The history of civil rights movements might be seen as small efforts to loosen parts of the knot. But, it seems, other parts of the knot tighten in response to progress. In the United States, the legal abolition of chattel slavery didn’t stop de facto slavery in the Reconstruction era. New legal, economic, and social arrangements quickly developed to justify new systems of unfree and forced labor. In the mid-20th century, as some thought leaders worked tirelessly to correct the popular fiction that black Americans were biologically inferior, white Americans adopted new fictions to justify their anti-black attitudes (Durrheim, 2014). When activists challenged these cultural explanations of inferiority, discriminatory practices found a new home in warped, “color-blind” conceptions of justice (Bonilla-Silva, 2006).

We should doubt there are any simple solutions to the systemic social problems. If activists fail to appreciate how parts of the system interact, their efforts will fail to secure significant or lasting progress. In some cases, their efforts have caused vicious backlash and regression. While the struggle for social progress may seem hopeless at times, individuals have tended to be optimistic about their ability to achieve moral progress. But we may overestimate their ability to do, be, or know better. We are, after all, products of the social systems we aim to change.
If we fail to understand the factors that get in the way of individuals’ abilities to change their behavior for the better, or to gain moral understanding, we may grow complacent with the illusory feeling we’ve done better and we may demonize others who need a helping hand. In this dissertation, my primary concern is with identifying obstacles to individual moral progress. I argue we might not be able to reduce our social biases and prejudices on our own; we may be rendered hopelessly ignorant by the cultures where we grow up; and there may be things we can't know in virtue of our social identities -- something we can’t control.

In Chapter 1, I consider the moral threat of unconscious psychological biases. These biases may seem like distinctly personal problems. But, I argue, we cannot simply attempt to fix the hidden threats “in our heads” the way we might treat a bad habit. We may be able to correct our prejudiced beliefs though personal reflection, careful attention, and self-guided activity. But this “self-help” strategy won’t help us address the unconscious, automatic ways we react to social information. I close by suggesting more promising strategies for reducing and eliminating these biases might target the social environmental conditions that sustain them.

In Chapter 2, I consider whether social practices can render us morally ignorant -- systematically insensitive to important moral considerations. I argue social practices can shape our belief systems, and distort our inquiry, in ways that leave us incapable of rationally correcting our ignorance. Because of this, some agents may not be morally responsible for their moral ignorance or for acting on that ignorance. Nonetheless, understanding how this ignorance is created and sustained may give us insights into better ways of engaging with agents who fail to understand they are participating in oppressive social practices.
Last, in Chapter 3, I consider how our social identities may compromise our judgments -- potentially threatening our own moral improvement, and obstructing understanding between social groups in conflict. Standpoint theorists once defended the claim that an agent’s social identity can give them an epistemic advantage when it comes to understanding social injustice and how to address it. I argue that attempts to explain this advantage in terms of some special power or capacity held by oppressed groups fail. Nonetheless, I argue we can recover the radical epistemological commitments of early Standpoint formulations by explaining this advantage in terms of the regular epistemic failings of socially privileged groups. Building on insights from the previous chapter, I present a novel account of epistemic differences between oppressed and socially privileged groups. The account defended explains why even well-intended agents with social privilege may be incapable of having reliable judgments about unjust social practices, their role in them, and how to ameliorate oppression.

While my arguments in these chapters may encourage a pessimistic view about individuals’ capacities for moral growth in societies plagued by unjust social practices, they can also offer hope. With a sober accounting of our limits as individuals, we can start to think more productively about what we can actually do to affect positive moral change in ourselves and in the world.

**Chapter 1 - Can We Fix Our Psychological Biases?**

1.1 - Introduction

As more Americans are confronted with the persistence of racial injustice, many are starting to consider whether they may unknowingly be part of the problem. While explanations
for phenomena like segregation, healthcare disparities, and mass incarceration involve complex interactions between history, politics, and law (Alexander, 2012; e.g., Cambria et al., 2018), many people are captivated by the idea that untreated, unconscious psychological forces play an important role in explaining persisting injustices. Though automatic patterns of thought, feeling, and action may only be part of the bigger picture, they implicate us in the problem in a very personal way. Few people are in a position to remedy hundreds of years of economic oppression, to prevent redlining, or to reform the criminal justice system. But most people believe they can control their thoughts, feelings, and actions.

If automatic, unconscious psychological mechanisms play an important role in the production and maintenance of unjust social conditions, these bad biases are morally significant. It seems we have some moral obligation to do something about the duplicitous moral threats “in our heads”, but it’s not obvious what we can or should be doing. One strategy for managing bad bias has gained enough popular attention to be deployed in training programs in universities, private corporations, and government organizations across the nation. This is the self-help strategy, which advises we treat bad bias like a bad habit. This strategy rests on the hypothesis that the biases relevant to the social reproduction of unjust conditions can be discovered, monitored, and ameliorated through a combination of personal reflection, careful attention, and self-guided activity. After employees attend a short educational workshop, attendees are expected to “get better” by a process entirely motivated and managed on their own.

The Unconscious Bias Training offered to faculty at the University of California San Francisco (UCSF)¹ is representative of many programs presently being run in universities and other entities across the nation. This training is modelled on peer-reviewed research that

---

¹ More about UCSF’s program can be found online: diversity.ucsf.edu/resources/unconscious-bias
prescribes using self-guided habit-breaking tactics on the “bad habits” of bad bias (i.e., Carnes et al., 2012; Devine et al., 2012; Forscher et al., 2017). Forscher and colleagues describe debiasing workshops as “based on the premise that unintentional bias is like a habit that can be broken with sufficient motivation, awareness, and effort” (Forscher et al., 2017, p. 133). For these researchers, the self-help strategy seems promising because it works in other contexts where people are struggling with the effects of powerful, unconscious motivation: nicotine addicts (e.g., Prochaska & DiClemente, 1983). According to this strategy, individuals can independently replace their bad habits over time with minimal training. Individuals first become aware of their undesirable habits and identify the events that trigger them. The smoker will identify the rituals that lead up to smoking, and might learn certain events and circumstances trigger their desire to smoke and related behavioral rituals. The individual then intervenes on the bad habit with a healthier one. The smoker might plan to practice new habits when they notice they are in the presence of a trigger. They may have to use a great deal of attention and effort to retrain their habits at first but, once the new habit is engrained, they’ve effectively eliminated their bad smoking habit.

Here’s how the habit-breaking self-help strategy for smoking cessation is adapted to combat bad bias. At bias training workshops, attendees learn about the nature of unconscious bias and take the Implicit Association Test online (at implicit.harvard.edu). They learn when implicit bias might be operative in the course of their professional decision making, then learn new habits they can do to prevent or override the “bad habits” of biased thinking, feeling, and acting. They are then expected to practice these new habits whenever they identify the triggers of bad bias. Attendees might be advised to use a counter-stereotype imagery technique. In contexts where they suspect bad bias might sabotage decision-making, they could attempt to ameliorate
the effects of bias by consciously imagining examples of people who don’t conform to negative race or gender stereotypes (Blair et al., 2001). Beyond the short (2.5 hr) workshops offered by these popular programs, attendees are not asked to come back for further training. The attendees are encouraged to deliberately practice their new debiasing tactics until it becomes a matter of habit. It is assumed, and expected, attendees will get better on their own.

The idea that undesirable patterns of thought, feeling, and behavior can be broken like a bad habit is not far-fetched. As Carnes and colleagues note, the self-help strategy is successful in the treatment of nicotine addiction. These strategies are also successful in the treatment of anxiety and depression. Cognitive Behavior Therapy (CBT) is often considered the standard for effective psychotherapies given its empirical track record (David et al., 2018) and prescribes a self-help model. The core idea of CBT is that anxiety and depression are the result of “warpy” thoughts: false beliefs and biased thinking. Depressed individuals may inappropriately interpret a minor setback as evidence for their incompetence, blame themselves for events that are beyond their control, or assume their attempts to succeed will fail despite having insufficient evidence to believe this (Parslow et al., 2006). While anxious and depressed individuals may feel they have little control over when these thoughts invade their thinking, CBT claims to offer control. Individuals are taught how to identify these bad habits of thought, to discover what triggers them, and to counter these biased judgments with new habits of more accurate thinking. While clinical studies typically have patients meet with a therapist for a number of sessions, some proponents of the therapy claim that patients can get better just through education and practice alone. (see Burns & Beck, 1999, pp. 29, 58–9).

If our biases are morally significant, we need to know how to manage them and their effects. If bad biases are simply bad habits of thought, the self-help strategy promises individuals
a method for taking some small effort to advance matters of social justice on their own. But there are several reasons to doubt the effectiveness of the self-help strategy. We should not believe that we can identify, correct, and monitor the biases that contribute to unjust social conditions through simple means of reflection, careful attention, and self-directed interventions.

First (§1.2), it’s clear some bad biases can’t be treated like bad habits. Most self-help programs are primarily concerned with implicit biases. I consider what we know about implicit biases and argue that the self-help strategy fails to be an effective means for managing implicit biases: agents can’t accurately detect and monitor their implicit biases in ways that would let them intervene. Moreover, while there are self-directed interventions which might limit the effects of implicit biases in important contexts like hiring decisions, we don’t yet know of self-directed interventions to mitigate implicit biases in everyday decision making.

Second (§1.3), the biases relevant to the production and maintenance of unjust social conditions aren’t exhausted by a consideration of explicit prejudices and implicit biases. It is a mistake for debiasing programs to neglect the influence of automatic cognitive and affective processes that aren’t simply described as “racist” or “sexist”. Some of these other bad biases may directly contribute to unjust conditions and some may interfere with the execution of the self-help strategy itself. Managing these biases may require a more robust solution than self-help. Though I cannot give an exhaustive consideration of these biases here, I focus my analysis on empathy gaps, identity-protective mechanisms, and some biases that arise from group conflict.

Last (§1.4), the self-help strategy implemented in popular programs also neglects the biasing effects of social and situational factors. Small environmental factors can have considerable influence over our thoughts and choices, and we may fail to detect their influence in ordinary life. I review some of the empirical findings demonstrating this influence. Situational
influence we believe we can resist may hamper our best intentions. These situational factors may influence us to behave in ways that directly contribute to unjust conditions, or they may interfere with the effectiveness of self-help strategies.

1.2 - The Problem with Implicit Biases

Popular debiasing programs do not assume that their half-day workshops are sufficient to eliminate bias in the workplace. Rather, they assume the work of debiasing is done by the agent after the workshop. The workshop trains the attendee to use a self-help strategy for correcting bad bias. Once attendees have learned about bias and ways to intervene, they are expected to carry out the work of identifying, correcting, and monitoring their bad biases on their own time. While we might rightly worry about whether people will be adequately motivated to correct their biases, the self-help strategy suffers from a deeper problem: we do not have sufficiently reliable access to our bad biases (or their effects) to monitor and correct them.

Reflection and attention are necessary parts of the smoker’s self-help strategy. Smokers can detect their urges and habits reliably enough to identify their habits, accurately correct them, and monitor how well their interventions are working (Prochaska et al., 1992; see Prochaska & DiClemente, 1983). Reflection and attention help smokers gain the information necessary to understand what conditions triggered their bad habits so they may effectively deploy some intervention. For example, smokers can observe they have unthinkingly started the behavioral ritual that ends in their smoking. With some additional directed attention and reflection, they can notice the circumstances in which their urges form. Then they can attempt to intervene. The smoker may be surprised to find they smoke more when they are bored or relaxed than when they are stressed. As they come to have a better understanding of these triggers, they may
Consciously attempt to practice healthier habits to replace the bad ones. Perhaps they try to take some deep breaths in these moments and distract their hands with a stress ball. Reflection and attention are also crucial in measuring their progress: they can observe their urges subside and the problem behaviors appear less frequently. If the intervention isn’t working, they can notice the lack of change in their habits, and attempt new interventions. Cognitive Behavioral Therapy also relies on this self-help strategy. If patients suffering from depression and anxiety could not access bad habits in their thinking, they could not identify the triggers of these thoughts; they would not be able to replace these bad habits with better habits of thinking or monitor their progress along the way.

Most adult Americans are able to report the contents of pervasive stereotypes that may bias their judgments: for example, whether a racial or ethnic group is stereotyped as competent or incompetent, socially warm or hostile (Abele et al., 2008). The self-help strategy may be useful to help individuals discover and manage their explicit prejudices insofar as they can reflect to find offending attitudes, use attention to identify when these attitudes are influencing their actions, and replace these bad attitudes with more desirable ones. But implicit biases -- their contents or their operation -- are not easily detected by careful reflection and attention. Despite the fact that popular debiasing programs are primarily concerned to combat implicit biases, the nature of these biases and the subtle effects they have make them poor candidates for self-help correction. I’ll characterize the difficulties of detecting these biases and their effects before considering ways in which the self-help strategy might still be relevant in the management of bad bias. I conclude that the self-help strategy fails to be an adequate way to manage implicit biases or their effects.
1.2.1 - Automatic, Unreflective, and Small

An action or judgment is *implicitly biased* when it is caused by automatic cognitive or affective processes that manage social information. The biases are “implicit” because they operate in ways that influence thought and behavior before the agent has any conscious (“explicit”) experience of judgment or choice. In this way, someone who is explicitly anti-racist might be surprised to find out they have implicit biases against non-white people.

In one experiment on implicit biases, Keith Payne (2001) and colleagues asked participants whether an ambiguous image of a tool was a weapon immediately after seeing the face of either a white or black person on a screen. This image was presented so quickly that participants could barely register what they saw (200 milliseconds). Nonetheless, participants were more likely to identify the ambiguous stimulus as a weapon after being shown a black face, and were much faster at correctly identifying weapons after being shown a black face (cf. B. K. Payne, 2001). One possible interpretation of these results is that humans are particularly good at encoding stereotypes from their environments -- regardless of their explicit endorsement or rejection of that stereotype. If a great deal of what we see in popular media shows black people as violent, we may begin to form a subconscious expectation. Thus, at some level of description, we expect to see a weapon after seeing a black face; however, our own reflection and attention on the matter may utterly fail to reveal this expectation. Implicit Association Tests such as these are carefully crafted to reveal the effects of automatic cognition that might be otherwise imperceptible.

The above study has been cited in explanations for why black civilians are disproportionately more likely to be the victims of police shootings (e.g., Engel et al., 2020). If implicit biases explain patterns of unjust social discrimination, they are correctly identified as
bad biases, and are the appropriate concern of debiasing programs. But some psychologists have questioned whether implicit biases are useful at predicting patterns of discrimination outside the lab. An early meta-analysis of research on implicit bias by Greenwald and colleagues (2009) claimed that measures of the bias were better predictors discriminatory behavior (e.g., reaction times, micro-behaviors, interpersonal interactions) than explicit measures of prejudice. In this way, they argued the Implicit Association Test was “predictively valid”. But, in competing meta-analyses, Oswald and colleagues were less hopeful that implicit bias was a useful predictor of discriminatory behavior outside the laboratory setting (2013, 2015). If implicit biases don’t explain patterns of discrimination in society, popular debiasing programs are miseducating attendees about the moral significance of implicit bias and, *prima facie*, fail to address any morally significant issue even if they prove effective at reducing implicit bias. But there is good reason to think implicit biases do explain patterns of discrimination and disadvantage outside the laboratory.

Meta-analyses considering the predictive validity of the Implicit Association Test agree that implicit bias predicts some discrimination, and that the measures of this effect are quite small (across studies, the correlational effect sizes ranged from 0.148-0.274). Where the competing meta-analyses disagreed was in their interpretation of this “small” effect. Greenwald and colleagues argued that smaller effect sizes have driven medical decision making; moreover, research has yet to account for the accumulated impact of such “small” acts of discrimination against a group over time (2015, pp. 558–559). Debiasing programs are rightly concerned with the moral significance of implicit bias; however, they seem overly optimistic about our abilities to detect implicit biases or their effects. Measures of implicit bias are better predictors of discriminatory behaviors than explicit, self-report measures -- having individuals reflect on their
attitudes. Moreover, the correlational effect sizes are small enough to spark academic debate. Individuals are unlikely to notice the ways in which their biases affect their behavior, and are far less likely to see these small effects as significant if social disparities are explained by the accumulation of these effects on another person over time. Implicit biases, and their effects, are not easily detected through careful reflection and directed attention.

But could the self-help strategy work even if implicit biases are not as easily detectable as other habits? If self-help works, controlled studies of habit-breaking interventions should have no trouble showing this. In comparison to a control group, workshop attendees should have reduced IAT scores after the workshop. Ideally, this effect would be sustained over time. Moreover, the differences in scores should be statistically significant -- i.e., we should be unable to explain it by chance alone. As yet, three controlled studies have tested the self-help strategy, each using a nearly identical workshop. Taken in total, they make a strong case against self-help.

In one controlled study, Devine and colleagues (2012) found a significant difference between IAT scores for attendees and the control four weeks after the workshop; however, this statistically significant difference disappeared when measures were taken eight weeks out. Forscher and colleagues (2017) attempted to replicate Devine’s study using three times as many participants (91 vs 292). They collected IAT scores every two days over two weeks and failed to find any significant differences between the IAT scores of attendees and the control. Carnes and colleagues’ (2015) controlled experiment was also a replication of Devine’s study. They tested an impressive 2290 people, and also failed to find any significant difference in IAT scores.

While these studies should have caused researchers to abandon faith in the self-help strategy, both Forscher (2017) and Carnes (2015) remained optimistic. In both studies,
researchers attempted to celebrate significant differences when measuring explicit states: participants’ beliefs that discrimination is wrong, or that they are biased (Forscher et al., 2017), and participant’s confidence in their ability to engage in equity-promoting behaviors (Carnes et al., 2015). But these self-report measures are not evidence attendees are any better at managing their bad biases. This measure, “self-efficacy”, is a gold standard for assessing habit-breaking interventions in addictive behaviors (cf. Prochaska et al., 1992). Except for the fact Carnes and colleagues want their data to support the self-help strategy, it is not clear why they think this is “strong evidence” their workshops are doing any debiasing. As demonstrated in the meta-analyses, IAT scores are far more reliable than any self-report metric at predicting discrimination.

The proponent of the self-help strategy might be able to modify their position. Even if we cannot simply reflect to identify whether we are implicitly biased, some argue individuals can still use self-directed activities to manage implicit biases or their effects. I consider trigger-avoidance strategies and counter-conditioning techniques. The first cannot be applied in most circumstances where the effects of bias should concern us, while the second is highly speculative given the current state of research.

1.2.2 - Avoiding Triggers

Some philosophers have been particularly hopeful that we can reduce the effects of implicit bias in some key, high-stakes contexts (such as unconscious discrimination in hiring or grading) regardless of whether we can detect their influence. We simply need to assume implicit biases will operate, and structure high-stakes decision-making environments to remove any conditions that could foreseeable trigger the bias (Madva, 2018; e.g., Washington & Kelly,
2016). For example, a hiring committee may attempt to remove information related to race or gender in the applications they review to avoid triggering implicit sexist or racist biases (Bertrand & Mullainathan, 2004).

In contexts where we are asked to evaluate individuals, if we can restrict our consideration to unbiased metrics, it’s reasonable to assume we can be unbiased. Symphony auditions performed from behind a screen can dramatically improve the chances of women getting hired (Goldin & Rouse, 2000), but there are not many other instances where we can “blindly” assess people in this way. Most hiring committees can only decide to interview someone on the basis of a resume scrubbed of identifying information. Then the task of avoiding exposure to information about the candidate’s social identities (like skin tone, gendered markers, or accents) seems nearly impossible.

Hiring decisions are certainly key opportunities where socially disadvantaged individuals could gain economic opportunity. Because of this, we should undoubtedly do everything we can to avoid allowing bad bias to affect the outcome. But, the morally significant effects of bad biases are not limited to the outcomes of high-stakes decision-making contexts. Ordinary interpersonal interaction seems just as important a target for intervention -- especially if implicit biases explain patterns of social disadvantage via the accumulation of small effects. Implicit biases could influence who we decide to engage in conversation, who we trust as a reliable informant, and who gets heard in contexts of group-based deliberation. They can influence what we end up believing and how we perceive situations (Saul, 2013). While effects in low-stakes contexts are far smaller than those that are observable in high-stakes contexts like job hiring, these small effects occur far more frequently in a person’s ordinary life. An approach that helps us avoid bias only in high-stakes situations neglects a significant part of the problem of bad bias.
The trigger-avoidance strategy is simply not feasible for managing implicit bias in ordinary interpersonal interactions, where information about others’ social identities is unavoidable. Even if it were theoretically possible to shield ourselves from this information, any practical implementation of the trigger-avoidance strategy in ordinary interaction would essentially prohibit ordinary social interaction as we know it.

1.2.3 - Counter-Conditioning

Another way to save the self-help strategy might involve finding ways individuals could implement interventions that have been found to be effective at reducing implicit bias. In theory, if research showed that an intervention was reliable at reducing implicit bias, and it could be practically implemented in the lives of most people, we could reliably ameliorate implicit biases without any need to detect and monitor them. Moreover, if the intervention directly mitigated the bias itself, this approach would not face the same limitations of the trigger-avoidance strategies: we could mitigate the sinister “small” effects of implicit bias by treating the problem at the root.

Considerable research has been done on a wide variety of inventive interventions but, sadly, extant meta-analyses struggle to draw strong conclusions about their effectiveness (Forscher et al., 2016). Nonetheless, Madva (2017) seems optimistic about the empirical record of counter-stereotype conditioning techniques (counterconditioning). Madva promotes these techniques as effective interventions individuals could easily implement in ordinary life. While most interventions focus on managing the effects of bias, or avoiding triggering bias, counterconditioning has been demonstrated to have a direct effect on implicit biases (Kawakami et al., 2000). In these studies, participants were shown pairs of faces and traits. When these pairings were consistent with known prejudiced stereotypes, participants were asked to report the
pair by pressing a key and thinking “No!” or “That’s wrong!” (Johnson et al., 2018). While original studies showed a reduction in bias after 480 of these “trials”, later studies showed promising effects with 200. These techniques could be practical if individuals could debias themselves with a 45-minute computer program simulating these trials, and a lasting reduction in bias could be achieved by doing this every couple of months (Madva, 2018, pp. 27–28).

Madva’s arguments might be attractive to proponents of the self-help strategy, yet we still don’t understand how to translate the insights of counterconditioning research into practical advice for individual treatment. Madva admits there is no direct evidence that these techniques reduce implicit bias for any more than a few days (2017, pp. 19–20). The promise of long-term interventions is generally complicated by research that shows implicit biases fluctuate over time and in context-dependent ways we’ve yet to fully understand (B. Keith Payne et al., 2017). Moreover, the 200-480 trials required for effective conditioning has only been demonstrated to mitigate implicit biases in a rather narrow way: for specific associations between a social group and a character trait. Research on implicit biases reveals that biases against one group involve many distinct associations: stimuli like black face, black voices, or black names have been found to provoke distinct cognitive, behavioral, and affective responses (Holroyd & Sweetman, 2016, pp. 87–89). It’s not clear if treating an association between black, masculine faces and weapons has any effect at reducing associations between black, female voices and assumptions about education. It’s doubtful whether negating the face/weapon association would have any effect on the discrimination other groups face on the basis of sexuality, religion, age, or disability. So the true practical burden of effective debiasing is unknown.

Further research into counterconditioning may reveal simpler and more effective ways to introduce counterconditioning in ordinary life. But whether it can be leveraged in a strategy to
help individuals reduce their own biases (and how) is yet unknown. One day people might “help themselves” with routine training exercises on a computer, but this is a far cry from the self-help strategies promoted in popular debiasing programs (cf. Forscher et al., 2017). Motivation, awareness, and effort are largely irrelevant if we can merely condition ourselves to be less biased.

1.3 - The Problem with Other Kinds of Biases

Popular debiasing programs assume that biases relevant to the social reproduction of unjust conditions (“bad biases”) can be discovered, monitored, and ameliorated by badly biased agents through a self-help strategy. I’ve argued that the self-help strategy is a not promising way to manage implicit biases -- the biases most commonly discussed in these programs. We do not have sufficiently reliable access to our implicit biases to monitor and correct them, and trigger-avoidance strategies do not account for a great deal of the damaging effects of implicit bias.

Another problem with the popular debiasing programs is that they are too narrowly focused on explicit prejudice and implicit bias. Cognitive or affective processes that operate on information about race and gender are obvious bad biases. Explicit prejudice involves false beliefs about a social group, and implicit biases associate a social group with negative qualities. But other kinds of biases can play a role in the production and maintenance of social injustice. Their relationship to racism and sexism is more complicated. They don’t always, or obviously, operate on information about categories like race and gender, but can have bad effects in virtue of the social contexts in which they operate.

Other biases that play a role in discrimination and advantage could include confirmation bias, attribution bias, self-serving bias, the availability heuristic, framing effects, the bandwagon
effect, clustering illusions, the empathy gap, the Dunning–Kruger effect, prestige bias, the halo
effect, mere exposure effect, status quo bias, social comparison bias, conformity bias, and
authority bias. Some might play a direct role in bringing about a bad effect, while others may
have more indirect effects by interfering with an agent’s ability to accurately reflect on, and
assess, their prejudiced behaviors. If attendees of self-help bias workshops are unaware of these
other biases, they could not reasonably be expected to use the self-help strategy to account for
their influence -- especially if these biases interfered with their ability to use the strategy.

In order to gain a better understanding of the biases that that play an important role in
production and maintenance of unjust social conditions -- like gender discrimination and racial
disparities -- we ought not limit our consideration to explicit prejudice and implicit social biases.
While accounting for every relevant bias would be a massive undertaking, I will only illustrate
the importance of this task by considering a few that seem particularly troubling: empathy gaps,
identity-protective mechanisms, and biases that arise from group conflict.

1.3.1 - Empathy Gap

One underappreciated bad bias might involve an “empathy gap”. For example, people
making judgments about their future behaviors in a “cold” state -- not influenced by visceral
drives like pain, hunger, or sexual desire -- underestimate the influence of these drives on future
and past decision making. For example, men may be more willing to engage in morally
questionable behavior when they are sexually aroused, and fail to account for this influence
when they predict their behaviors from cold states (Ariely & Loewenstein, 2006). Thus, men
may be ineffective at preparing for what they will do “in the heat of the moment”, or may fail to
understand how they were influenced by visceral states in the past. A well-intentioned man
might learn about forms of sexual harassment and learn, from women colleagues, that he has a tendency to make them uncomfortable with sexually-charged comments and jokes. In a cold state, he may fail to realize that these events are occurring when he is in a state of arousal. He underestimates the significance of the bodily cues that make the comments so hostile: excessive staring, facial expressions, and subtle invasions of personal space. He plans to address his behavior by simply making a list of topics he shouldn’t discuss, but fails to account for the bodily effects of arousal. While he mostly succeeds in avoiding sexual topics, his colleagues struggle to explain how he’s now managed to make normal professional conversations feel sexually charged. On reflection, he knows he’s been very careful with his words. He’s even made a rule not to touch anyone in the office. Lacking insight into how his arousal influences his behavior, he is satisfied with his efforts and decides his colleagues’ concerns have no merit. In theory, he could figure out how to manage his behavior. In practice, Human Resources’s misguided bias workshop gives him a false sense of confidence and makes it more difficult for him to gain clear feedback about his behavior.

While empathy gaps might hamper individuals’ ability to evaluate their own decision making, they may also hamper individuals’ ability to understand others in contexts of social conflict (e.g., Handgraaf et al., 2008). Conflict over matters of sexism or racism is often “heated” in the sense that one group’s claims on another are sourced in physical and emotional pain. If white people underestimate black people’s experience of pain (Hoffman et al., 2016), even well-intentioned anti-racist white people may systematically struggle to grasp the significance of the moral demands being made.

It’s worth mentioning the use of the self-help strategy for smoking cessation has been criticized for failing to account for empathy gaps (Lowenstein, 1999, p. 240). In theory, self-help
strategies could be enhanced with special attention to the empathy gap (Lowenstein, 1999, p. 257). But we have a decent empirical understanding of the role of visceral drives in addiction. It may be considerably more difficult to fully understand how the empathy gap contributes to sexism or racism, especially since there has not been much research on the matter (Gair, 2017; Gracia et al., 2014; cf. Gutsell & Inzlicht, 2012; Seager & Barry, 2019). In practice, our lack of actionable insight means self-help bias workshops are unlikely to adequately prepare attendees to account for empathy gap.

1.3.2 - Identity and Group Conflict

Another set of underappreciated bad biases involves how we are influenced by social groups that inform our identity -- especially those groups that aren’t explicitly organized around issues of race and gender. Research has found that surprisingly arbitrary social groupings can provide the basis for biased thinking and behavior (e.g., Tajfel, 1978). Social groups that inform our sense of identity can bias our perceptions of others, the way we process information, and how we conceptualize social justice (B. Keith Payne & Cameron, 2010, pp. 8–11; Sherman & Cohen, 2002; Van Bavel et al., 2014). While these groups can have considerable influence over what we believe, we seem to be particularly bad at recognizing this influence (Cohen, 2003; Mackie & Queller, 2000). An agent may believe they are defending their neighbors, national identity, or political party without realizing the practices or policies they defend non-accidentally reinforce racial disparities (e.g., Sears et al., 1979; Shen & LaBouff, 2016).

If an agent holds a prejudice in virtue of some false belief they share with a group, the prejudice biases cognitive and affective processes due to the content of the belief. Because beliefs are mental states that are susceptible to rational correction, we might reasonably think
some process of self-help can be effective at addressing these biases. Traditional diversity and sensitivity training typically aim, on some level of description, to help agents identify and correct biasing beliefs. Perhaps it’s reasonable to expect that agents trained to rationally engage with prejudiced beliefs about race or gender could discover prejudice in other types of beliefs. But the self-help strategy can’t help an agent get rid of a prejudice if the agent can’t recognize it as a prejudice. Social identities can interfere with sober recognition. For example, being discreet about one’s religion is part of what it means to be French. French national identity involves political and social commitments to a notion of secularism: the principle of laïcité. A French person might be pro-immigration and be against religious discrimination, while maintaining negative attitudes towards Muslim women and anyone else who broadcasts their religion. In some way, a woman who is observing Islamic rules of modest dress (hijab) is interpreted as anti-French, and an insult to French values (Winkler, 2016). Despite condemnation from around the globe, and much internal debate, many French people maintain certain forms of social and legal discrimination are justified by laïcité. Getting rid of this prejudice is not simply a matter of discovering some belief that is obviously wrong and changing it. Since racial attitudes in the U.S. are also shaped by complex values systems and ideas (Bonilla-Silva, 2006; C. W. Mills, 2014), white Americans face similar struggles in recognizing their bad biases.

While an agent’s bad biases might be caused by beliefs that are important to a social identity, an agent’s identity can also bias information processing when the agent perceives the group’s beliefs and values are challenged in group conflict. These biases can further complicate an agent’s ability to reflect, monitor, and correct their biases. When we face threat to the group’s goals or values, we become more motivated to search for information that favors the group’s beliefs (Schulz-Hardt et al., 2000), and to interpret challenging evidence in non-threatening
ways (Lord et al., 1979; Sherman & Cohen, 2002). When confronted with opposing views, we may be motivated to misconstrue the views and testimony of the other side, and to overestimate their biases, while gaining more confidence in the objectivity of our own views (Nyhan & Reifler, 2010; Robinson et al., 1995; Sherman et al., 2003). Even if debiasing workshops could address the complicated ways in which our social identities can interfere with the self-help strategy, merely educating attendees could result in them seeing the training as a threat and rejecting the training. Ameliorating these kinds of biases may require a more gentle approach.

We must look beyond explicit prejudice and implicit social biases if our aim is to understand the role biased thought, affect, and action play in the production and maintenance of unjust social conditions. The prevalence of unexpected, unreflective, automatic affect and cognition seems to pose a serious threat to our ability to assess our behavior (Pasnau, 2020; e.g., Saul, 2013). In some cases, we may have strong reasons to doubt our ability to have much control over these processes (e.g., Bargh, 1999). The proponent of the self-help strategy may attempt to continue to add new “modules” to the workshop to address every new charge they fail to account for bias. But this ad-hoc approach isn’t enough to defend the efficacy of the self-help strategy. Some of these biases -- especially those arising from identity and group conflict -- pose a threat to the efficacy of self-help itself. Ameliorating bad bias isn’t as simple as identifying a group of behaviors that makes a bad habit and replacing those behaviors. In some cases, coming to terms with our bad “habits” of bias might involve a much more personal therapeutic approach that includes learning how to empathize with others and interrogating who we believe we are.
1.4 - The Problem of Social and Situational Influence

Popular debiasing programs assume that biases relevant to the social reproduction of unjust conditions ("bad biases") can be discovered, monitored, and ameliorated by badly biased agents through a self-help strategy. I’ve argued this self-help strategy isn’t promising for managing implicit biases -- the “star” threat in these workshops. Moreover, I’ve argued the self-help strategy doesn’t seem like a promising intervention in light of other biases and prejudices that are relevant to the production and maintenance of unjust social conditions. I close with one more criticism of the self-help strategy: the situation in which biased behavior occurs may be just as important in our efforts to reduce bad bias.

A large number of studies by psychologists (see also J. M. Doris, 2002; e.g., Ross & Nisbett, 2011) and behavioral economists (Ariely, 2013; e.g., Thaler & Sunstein, 2008) have demonstrated that small environmental manipulations can have surprising effects on our decision-making and action. In many cases, though the experimental manipulation is the best predictor of what people will do, subjects in the experiment often believe they are making the decisions themselves. While little of this body of research attends to matters of racism or sexism, there has been a great deal of attention on moral behaviors.

Early study into situational effects on human judgment involved asking individuals to do a simple task comparing the lengths of lines. Without social pressure, subjects made mistakes less than 1% of the time. But, when they were making judgments as part of a group, they would increasingly report incorrect answers (36% in the original). The other subjects in these groups were actors who were instructed to give the same, but incorrect answer. Solomon Asch and colleagues were attempting to determine how the known opinions of a majority group could
influence a person’s beliefs (E. Asch, 1951; 1956). It appeared that subjects were willing to publicly agree with majority opinion despite their better judgment.

Infamous studies like Milgram’s shock experiment (S. Milgram, 1963) and Zimbardo’s prison experiment (Zimbardo et al., 1971) attempted to see how susceptible our moral thought and behavior were to social influence. There’s much debate around the best ways to interpret these experiments (Le Texier, 2019; see Stanley Milgram, 2017). Nonetheless, both revealed that situational factors could pressure people to commit wrongful acts despite their better moral judgment. In Milgram’s initial experiments, subjects delivered shocks of increasing voltage to, what they believed was, another subject in the experiment -- the “learner”. Under the verbal instruction of a presumed authority figure -- the experimenter in a lab coat -- all subjects continued to deliver shocks despite hearing the learner beg them to stop, and continued until the learner became unresponsive (“300 volts”). A considerable number of subjects (65%) continued to deliver another ten shocks (up to “450 volts”). This study has been replicated many times, and with slight changes to the set-up to determine the strength of the influence of the experimenter. Consistently, no fewer than 28% percent of subjects carry out the full series of shocks when directed to do so (Blass, 1999). In debriefings with subjects, and in observations of their behavior, it is clear they suspected they were committing wrongful acts (Stanley Milgram, 2017; esp. Sheridan & King, 1972). In Milgram’s replications, the few subjects who refused to obey instruction didn’t attempt to have the experiment shut down, and they only attempted to assist the harmed learner after seeking permission from the experimenter (WGBH, 2001, sec. 11:25:12).

On any interpretation of these experiments, agents were predictably influenced to act against their better moral judgments due to an outside influence. Related studies on situational effects suggest we are more inclined to commit wrongdoing when we see others around us doing
it (Gino et al., 2009); similarly, we may be less inclined to act on evidence that someone needs help if people around us appear indifferent to that evidence (Darley & Latané, 1968). We might also abandon our moral principles when we are in a hurry (Batson & Darley, 1973). Some studies on situational influences show the situation can encourage us to act on our moral commitments. For example, we may engage in more moral behavior if we have the sense that others are watching us (Bateson et al., 2006, 2013) or if we have evidence that others around us are doing the right thing (Goldstein et al., 2008; Nolan et al., 2008). While we may not always be obligated to help others or to be charitable, we may be more inclined to act in these morally valuable ways if the room we’re in smells nice -- like fresh baked cookies or coffee (Baron, 1997; Liljenquist et al., 2010). This second set of findings might give us optimism about our ability to promote more moral behavior; however, they still give us reason to doubt just how much conscious control we have over what we do.

Due to questionable research practices and biases in what gets published, there has been considerable concern over whether some of the most interesting findings in psychology can be replicated (e.g., Klein et al., 2014; Open Science Collaboration, 2015; Wiggins & Christopherson, 2019). We should expect that the conclusions of any one of the studies above, taken in isolation, overstate the power of any particular situational influence (J. Doris, 2015, pp. 44–9). But the hypothesis that our judgments and behaviors are surprisingly influenced by small environmental factors is supported by converging evidence from a great number of studies. In an academic apocalyptic scenario in which (say) only 30% of these studies could be replicated, we would still have considerable reason to think situational factors pose a serious threat to our abilities to act and judge in accordance with the moral beliefs we endorse.
Situational influences bias human thought and behavior, and these biases can play a role in unjust social outcomes. In a workplace, many of the men might regularly get away with subtle sexual harassment and have some social incentives to do this. A man committed to not sexually harassing women might start to see some subtle forms of harassment as permissible, or give under the social pressure and act in sexist ways if he thinks he will impress the boss. White neighbors in the suburbs may appear to be unconcerned about the moral claims being made by black protesters in the city. A neighbor committed to racial equality may fail to look further into these claims, or take on negative and dismissive attitudes toward the protest if all her neighbors are voicing these attitudes. These situational forces can undermine the self-help strategy insofar as agents can be unwittingly influenced not to apply the strategy in certain circumstances: distorting an agent’s ability to acknowledge their biases, or to recognize the full significance of their biases, and skewing their ability to monitor how successful they are at reducing bias (see Noon, 2018 for a more comprehensive consideration). More generally, agents might abandon their training if they don’t believe others around them are applying it, or if others around them hold negative attitudes about the training. It would not be surprising if employees more frequently shared negative attitudes about mandatory training, and if the prevalence of post-workshop grumblings disinclined attendees from putting their knowledge into practice.

Given the complexity, variety, and unpredictability of most situational influence, it’s unclear whether self-help strategies could be effectively used to combat them. We may be able to use available research to self-direct in ways to account for the biasing effects of situational influence. But the available research isn’t an exhaustive guide of biasing influence and, due to the large number of potential threats to account for, attempts to manage all known influences might inhibit our abilities to carry out the various other responsibilities we have. There are only
so many hours in a day, and most people might be inclined to abandon the self-help strategy altogether considering the burden.

When a study on situational influence gains popular attention, popular media outlets tend to publish “commonsense” advice about combating the new threat. As yet, little (if any) of this commonsense advice has been tested empirically. Many programs aimed at reducing sexual violence employ a self-help strategy is to address our tendency to not help or intervene on cases of wrongdoing when others are present. Attendees at workshops are educated about this bystander effect, how to identify situations where it happens, and are trained in ways to replace their habit of inaction with effective strategies to intervene on sexual harassment and sexual assault. Sophisticated programs may even try to thoroughly train attendees in simulated scenarios. But recent study has been concerned with determining if this self-help strategy makes a difference. In follow-ups with attendees of these workshops, researchers find poor evidence that attendees are able to use self-help to combat the bystander effect (e.g., Woods et al., 2020). A meta-analysis of 151 of these studies argues that there are some workshops that appear to promote bystander intervention, but the success of these workshops can not be explained by attendees' use of the self-help model (Kettrey & Marx, 2020).

Proponents of self-help might strategically ignore situational effects because they appear to be more generic threats to our ability to carry out our good intentions. But situational effects may have a more direct relationship on the prejudices and implicit biases that are targeted in self-help debiasing workshops. The strength of our implicit biases may correlate over time with our exposure to racist or sexist representations in our environment (B. Keith Payne et al., 2017). There is more robust empirical support showing that our explicit prejudices (and other negative social attitudes) are acquired or made stronger by racism and sexism in our environment (Jolley
et al., 2020; Shaver et al., 2017; e.g., Soral et al., 2018; Tropp, 2003). If agents’ biases are re-learned, reinforced, and restored by the prevalence of racism or sexism in their environments, any positive effect of self-help strategies may be immediately undone over a matter of days. We encounter racism and sexism in popular media representations, but also in our ordinary interpersonal interactions. A greater threat may be information and entertainment services that deliver content in response to how humans use it. The algorithms behind Google search, Facebook, and Twitter learn our biases and reinforce them, and may make our exposure to re-learning triggers unpredictable and unavoidable (Noble, 2018).

The problem of biases being re-acquired through environmental exposure casts doubt on the effectiveness of individualistic approaches to debiasing. The self-help model may fail to sustain debiasing efforts in a world plagued by similar biases. Ameliorative efforts might make a greater impact on our bad biases by targeting the environmental factors that cause us to learn and relearn bad bias (Haslanger, 2015b; e.g., Huebner, 2016). Individuals may be able to limit their exposure to a limited extent but, depending on the context, efforts to eliminate exposure might require agents to shut themselves off from the world, the most effective strategy for reducing the negative impact of our own bad biases. This reclusive strategy misses the point of debiasing: securing widespread positive social outcomes. A concerted effort to reduce prejudiced representations in a society, and the practices that sustain them, could reduce biases in large populations of people without those people even trying. This effect might be enhanced with a parallel effort to promote more anti-racist and anti-sexist representations. This might amount to a counter-conditioning bias intervention on a massive scale. But it is considerably difficult for those who lead social justice efforts to accomplish this kind of widespread change. There are moral and political ramifications to this kind of intervention that might make some solutions
overall morally undesirable. Moreover, it has been demonstrated that some anti-prejudice messages can backfire: increasing prejudice and bias in the intended audience (Legault et al., 2011). Since we are yet unsure about the effectiveness of large-scale environmental interventions, or what kinds of interventions are most effective, the practical costs and difficulties of pulling it off doesn’t make it a promising alternative to self-help strategies. Putting on a short workshop is certainly an easier way to gain the sense one has made moral change in the world, even if it fails to affect any lasting change. Nonetheless, environmental interventions at smaller scales have lower stakes, and may complement existing efforts to help individuals manage their biases. Since people spend a considerable amount of time at school and work, interventions on these environments might support positive debiasing effects that are sustained beyond school and work. Despite the complexities of implementing environmental interventions, the value of small-scale efforts are certainly worth further empirical investigation.

Situational influences are among the automatic cognitive and affective threats that may explain unjust social outcomes. If our aim is to reduce the negative moral impact of biases in our thought, affect, and action, we also need to consider situational influence. Like the biases considered in the previous section (§1.3) situational factors may influence us to behave in ways that directly contribute to unjust conditions, and they may threaten the effectiveness of self-help strategies.

1.5 - Conclusion

The self-help strategy for bad bias justifies what is taught in the debiasing programs popular among universities, governments, and large corporations across the contemporary United States. I’ve argued we should doubt the self-help strategy for bad biases. The strategy fails to
have demonstrated effectiveness on implicit biases (§1.2) and, in practice, the strategy has failed to account for other biases that are involved in the social reproduction of unjust conditions (§1.3) or situational influences (§1.4).

Very few of the biases relevant to the social reproduction of unjust conditions can be discovered, monitored, and ameliorated through a personal reflection, careful attention, and self-guided activity. Many bad biases are difficult to detect and monitor, and existing research doesn’t provide support for simple, practical interventions individuals can use in their daily lives. Some interventions appear to be effective in very limited contexts. More ambitious interventions may be overwhelmingly impractical, and study on practical interventions has yet to reveal individuals are capable of making lasting change to their bad biases. The general effectiveness of self-help strategies may be hampered by other biases (§1.3) and situational influence (§1.4). It is practically impossible for individuals to attempt to address all their bad biases on their own.

I’ve suggested that strategies for ameliorating bias through changes to the social environment may compliment self-help efforts, even if they are not outright considerably more effective. In the next two chapters, I continue to develop arguments that individuals’ attempts at moral improvement may be thwarted by the social environment: social practices can render us helplessly morally ignorant, and may prevent some agents from understanding the nature of oppressive social practices.
2.1 - Introduction

When an agent’s ability to gain knowledge or understanding is obstructed by the social practices of their community, that agent is *structurally ignorant*. Ancient Greek slave owners, sexists in the early 20th century, and contemporary consumers all appear to have been confused about some of the important moral issues of their day. In each case, pervasive cultural meanings, values, and social practices may explain this widespread ignorance. When we are engaged in a social practice, our thoughts, feelings, and actions are guided by culturally shared concepts, beliefs, and other attitudes (Haslanger, 2015a). In some cases, social practices may systematically distort individuals’ understanding of the moral significance of their actions. They may believe that they are acting in morally permissible ways when, in fact, they are engaged in moral wrongdoing.

Some have argued that culture is a poor excuse for moral ignorance. Moody-Adams argues social practices can never render agents unable to know right from wrong (Moody-Adams, 1994, p. 293). She claims that agents are always capable of questioning their social practices (Moody-Adams 1994, 296 fn14). The so-called structurally ignorant merely avoid questioning their social practices because they are unwilling to accept the moral truth such scrutiny could reveal (1999, p. 183). She argues they suffer from willful (or “affected”) ignorance and are blameworthy for their ignorance and their ignorant acts. On the other hand, we may feel hesitation to blame agents since they are products of their culture. Miranda Fricker defends the idea that “judgments of blame are out of order when people’s behaviour is the
product of ethical thinking that was routine in their ethical culture” (2007, p. 106). She agrees that some agents can break away from prevailing social mores, but argues this requires “exceptionally imaginative” moves of thought. These extraordinary thinkers are key to a society’s moral progress and warrant the thought that ignorant agents could do better. Nonetheless, Fricker maintains we lack justification for blaming them.

With few exceptions, most accounts of moral responsibility have an “epistemic condition”: in order for an agent to be fully responsible for some action, the agent must have some awareness of the moral significance of their action and its consequences. While there is considerable disagreement about the exact details of this awareness, most accounts agree that an agent is blameworthy for acting on their ignorance only if they are blameworthy for their ignorance. (Wieland & Robichaud, 2017, pp. 1–28). Elsewhere, Fricker suggests we may blamelessly inherit biased epistemic practices from our culture (2016). Learned patterns of judgment can hide their epistemic significance from us, and we may fail to realize this without being culpably negligent.

Our attributions of blame and responsibility are morally significant. If we have a moral duty to help others achieve moral understanding, we may be required to intervene on structural ignorance. We can only do so effectively if we accurately understand how social practices can render populations of people insensitive to the truth. If some structurally ignorant agents are not to blame, publicly blaming them and demanding that they “try harder” may fail to help them. In some cases, this derision can backfire -- driving the ignorant to be even more deeply committed to their confused moral outlook. Insofar as attributions of blame and responsibility guide our reactions to wrongdoing, we may actually wrong the accused with inaccurate judgments. Our impulse to punish wrongdoing may fail to track factors that matter when establishing
responsibility -- such as the agent’s intentions, their ability to foresee outcomes, or their ability to control their action (Mazzocco et al., 2004; M. E. Oswald & Stucki, 2009). If we actually care about wrongdoing, we ought to be concerned about being accurate about our attributions of blame and responsibility.

Could social practices render us insensitive to moral facts? If so, are we always blameworthy for structural ignorance? In this chapter, I argue that there are cases where social practices render agents morally ignorant and where agents are not responsible for their ignorance (§2.2). Shared social practices can rationally justify false moral beliefs and, further, practices of asking questions, trusting, and interpreting information can render agents unable to rationally correct their ignorance. Agents can’t be responsible for failing to gain moral understanding if they can not rationally correct their ignorance.

I then consider three objections to this view. Isn’t moral knowledge easy? (§2.3.1); aren’t these agents just willfully ignorant? (§2.3.2); and aren’t they blameworthy regardless of their ignorance? (§2.3.3). The first two are objections found in Moody-Adams’ arguments, while the last is found in Elizabeth Harman’s development of Moody-Adams’ position. I conclude none of these succeed at showing social practices can’t render an agent morally ignorant, or that these agents are always blameworthy for this ignorance. We may largely be guilty for failing to question social practices, and for being willfully ignorant, but there may nonetheless be important cases where social practices interfere with moral understanding.

2.2 - Structural Ignorance Can Excuse Wrongdoing

Megan Phelps-Roper was raised by the cult-like Westboro Baptist Church in Topeka, Kansas. Born into a radical system of moral and religious beliefs, Megan learned her purpose
was to carry out God’s will in a holy war between good and evil. She understood\(^2\) she had a duty to inform sinners about their immoral acts, giving them the chance to correct their wrongdoing. She would regularly join the Church in their ministry, which amounted to daily protests at funerals and other sites where people mourned tragic events.

Part of her ministry involved insulting the grieving family of fallen soldiers: using disrespectful language to inform them that their child deserved to die as a punishment for their sins. The fact her insults inflicted emotional harm on the grieving was clear to Megan, but she believed this was the only effective means to reach sinners. These “little” cruelties were necessary to give sinners the chance to repent and avoid the eternal pain of damnation; thus, they were morally justified. These false beliefs explained Megan’s *moral ignorance* -- why she failed to understand that insulting grieving families is morally wrong. While her corrupt moral outlook was the result of intense indoctrination from a young age. She was taught practices of thinking, feeling, and behaving that prevented her from truly questioning the Church. Such practices dictated which questions could be asked and how they could be answered, who could be trusted and who should be distrusted, and how evidence that might cause her doubt could be explained away.

Megan hurt others because she falsely believed her action was morally justified. She would be responsible for her wrongdoing if she were responsible for her false moral beliefs. But her ignorance appears to be the result of “brainwashing”. Might she be excused for her ignorance and her wrongdoing? Or is she responsible despite the fact she was brainwashed?

---

\(^2\) In the articulation of this case, epistemic success conditions like “understand” and “knowledge” are not used factually, but as a description of what Megan took herself to know or understand.
2.2.1 - Responsibility for Wrongdoing

Our judgments about moral responsibility retrospectively evaluate whether an agent’s behavior issued from their exercise of certain powers and capacities. This sense of responsibility is not synonymous with “duty” or “obligation” as an agent can be excused for failing to act on their obligations. The actions of a responsible agent can be attributed to them in a way that informs our moral assessments of them; moreover, agents who are responsible can be held accountable for their failure to meet some moral standard (e.g., G. Watson, 1996). There is little consensus among philosophers regarding the nature of moral responsibility. A thorough consideration of all the important debates would require its own book. Nonetheless, most accounts agree that two conditions must be met for an agent to be responsible: (1) their action must be the product of a sufficiently free choice, and (2) they must be sufficiently aware of the moral significance of the action and its consequences. These can be called the freedom condition and the knowledge condition, respectively. What follows are two categories of excuses agents might have for conduct that fails to meet moral standards: a force excuse if the freedom condition was not fully satisfied, or an ignorance excuse if the knowledge condition was not fully satisfied. An excuse does not show that the subject was not actually subject to the relevant moral demand, but describes a state of affairs that renders the agent less than fully responsible for their moral failure. In some cases, excused agents may be blameless for their wrongdoing. Much of the disagreement in the responsibility literature centers around the precise characterization of the freedom and knowledge conditions.

Megan Phelps-Roper might not be responsible for her actions if she didn’t freely choose to do them. But Megan recounts choosing to insult the grieving strangers as an adult: considering
whether the protests were morally permissible, and concluding they were, before joining in. If Megan faced social pressure to go along with the crowd, it does not seem like a strong excuse. Any social exclusion she may have faced for not protesting was minor. We know this because she decided against joining Church protests for at least a year while remaining a member in good standing (Chen, 2015). Thus, the social pressures she faced to act were fairly mundane. College fraternities regularly use stronger social sanctions to get their pledges to perform more severe acts of harm. Even if we accept that such social pressure has some exculpatory force, it does not seem strong enough to render them blameless for their action. Moreover, this kind of social pressure is ubiquitous. It is not an excuse that shows Megan faced any special pressure, and it fails to account for Megan’s brainwashing.

Megan freely chose to engage in wrongdoing because she falsely believed her actions were morally justified. Her corrupt moral outlook might provide an ignorance excuse: she failed to understand the moral significance of her actions. But ignorance does not always provide a good excuse. Agents who avoid certain forms of inquiry because they suspect it would result in some inconvenient moral truth seem to be fully blameworthy both for their ignorance and for acting on their ignorance. Ignorance fails to be an excuse for an agent when they are responsible for some benighting act, an act or omission that explains their ignorance (H. Smith, 1983). If Megan is excused for her wrongdoing, it’s because she isn’t responsible for her ignorance.

2.2.2 - The “Brainwashing” Excuse

At first glance, explaining why Megan Phelps-Roper isn’t responsible for adopting the beliefs of the cult-like Westboro Baptist Church seems straightforward if we can correctly claim she was “brainwashed”. It is popularly thought that cults use overt methods of psychological
control -- like deception and coercion -- to indoctrinate and maintain faithful believers. In some recorded cases, cults and other extremist groups use blackmail and physical threats to recruit new members and to prevent existing members from having contact with the outside world (Hassan, 1990, 2012). If agents form their beliefs because of coercion like this, we have compelling reasons to believe they are not responsible for their beliefs. They have a clear force excuse: their thoughts, choices, and actions are not performed freely. But the view that faithful believers are the product of psychological control tactics has been exposed as largely pseudo-scientific (American Psychological Association, 1987).

In most groups we might label cults, member’s thoughts, feelings, and behaviors are regulated through their voluntary participation in shared social practices (Hassan, 1990; cf. Zablocki, 2001). Versions of these social practices can be commonly found in more mundane social groups; however, cult practices tend to have peculiar epistemic consequences. These practices typically operate in ways to suppress certain forms of inquiry, to silence or shun testifiers who defy the cult, and to skew how members interact with outside information. In effect, they immunize the cult from criticism and other external threats. Social epistemologists have argued that these practices, and the epistemic environment they create, can establish conditions where agents can adopt extreme beliefs rationally (Hardin, 2002; Sunstein & Vermeule, 2009).

Megan Phelps-Roper’s moral ignorance is the product of the practices of reasoning and inquiry the Church taught her. As soon as Megan could talk, she learned the values, beliefs, and practices of the Church from her parents. Her parents taught her to trust and obey Church elders because, they explained, the elders most reliably understood God’s will. Her parents and the
elders taught her to actively distrust anyone who didn’t belong to the Church, even if they were nice to her. Megan was not discouraged from critically questioning Church doctrine. In fact, Church elders would encourage the debate. These would prove opportunities to teach her how to think. She would learn how to interpret the Bible, the law, and current events to justify their worldview. She would also learn how to respond to people who disagreed with her views and how to manage outside information that challenged her beliefs. She became so adept at explaining and defending Church doctrine that she represented the Church in media interviews before graduating sixth grade.

Some cults strictly prohibit members from talking to non-members, or from accessing information from the world outside the group. If Megan’s free access to information was constrained by such factors, these factors might easily be leveraged in a force excuse. But Megan’s ministry required her to convince outsiders that they were living immoral lives. In practice, Megan spent a considerable amount of her time carefully explaining and defending her beliefs to hostile outsiders. This fact may test our intuitions about her responsibility: it seems she was actively participating in the maintenance of her ignorance in the face of overwhelming evidence that she was wrong. Nonetheless, I’ll argue that Megan isn’t fully responsible for her ignorance because she was meta-ignorant: she was ignorant about the epistemic significance of the practices that maintained her ignorance.

2.2.3 - Responsibility for Ignorance

If we are responsible for our moral beliefs, it is not because we are responsible for their formation. Largely, our beliefs are formed involuntarily in response to reasons or available
evidence. Because we have little to no direct voluntary control over the formation of our beliefs, we fail to meet a key condition that would make us responsible for acts of believing (W. P. Alston, 1988; Audi, 2001; Nottelmann, 2007, Chapter 8). Nonetheless, we have voluntary control over many actions that indirectly affect what we believe. We can choose to question our beliefs, how we gather evidence, how we manage that evidence, and how carefully we reason to some conclusion. Thus, if we are responsible for our ignorance it is because we are responsible for actions or omissions that indirectly affect what we believe; that is, we are responsible for some benighting act (H. Smith, 1983).

Megan’s ignorance is explained by the practices of reasoning and inquiry the Church taught her: acts of questioning or failing to question, how to assess testifiers and testimony, and how to manage and evaluate evidence. While her ministry gave her opportunities to correct her false moral beliefs as an adult, Megan remained steadfast in her beliefs. In the absence of coercion, it’s tempting to think she must be making some kind of epistemic error. Perhaps she routinely failed to update her beliefs in light of the evidence, failed to notice her logical errors, or consistently resorted to informal fallacies in her debates. These mistakes might give us reason to blame her -- especially if she avoided careful thought and inquiry out of a desire to avoid inconvenient truths about herself. But Megan is not guilty of skirting any epistemic obligation.

Broadly, epistemic obligations are obligations to act (or refrain from acting) in order to achieve epistemic success -- like truth, justification, or knowledge (Feldman, 1988, p. 376). She only trusted testimony from sources she believed to be trustworthy and reliable, and was in the habit of thoroughly questioning what she was told. Megan voraciously sought out new sources of information and made genuine attempts to understand alternative points of view. Some Christian parents control their children’s access to various media in order to shelter them from sin, but
Megan was free to watch any television she wanted, listen to any music she wanted, and had unrestricted internet access. These freedoms were all intended to promote her ministry: to understand the sinful nature of popular culture, and to reach sinners over social media. Fred Phelps, Megan’s grandfather and founder of the Church, was a lawyer who intently trained his children and grandchildren to avoid logical errors and fallacious forms of reasoning. Megan was not only careful in her thinking, but she routinely considered the rational basis for her beliefs -- as anyone might if their beliefs were routinely challenged. Megan might be said to have been epistemically virtuous; nonetheless, she seemed to be trapped, unable to correct her ignorance.

Megan’s beliefs would prove intractable because they were formed in an echo chamber. Many of us may form our views in communities where certain testifiers are excluded. This can occur through ordinary processes of community formation and, in some cases, this “epistemic bubble” could provide overall epistemic benefits to a group (e.g., in scientific communities). While epistemic exclusion can certainly lead a group to gain confidence in false beliefs, members in a bubble are still open to rational correction from outside voices. By contrast, echo chambers isolate their members from outside influence. As C. Thi Nguyen explains, an echo chamber is

an epistemic community [that] creates a significant disparity in trust between members and non-members… by excluding non-members through epistemic discrediting, while simultaneously amplifying members’ epistemic credentials… [N]on-members are not simply omitted or not heard, but are actively assigned some epistemic demerit, such as unreliability, epistemic maliciousness, or dishonesty. (Nguyen, 2018, p. 6)

40
The credentialing practices of an echo chamber insulate members’ beliefs from rational correction by outsiders. Many of the practices Megan learned in the Church had this effect. She was taught non-members were dishonest and unreliable before ever being exposed to their views. Though Megan would carefully assess new evidence and the testimony of outsiders, they could only be rationally evaluated in light of the beliefs she already held. Among these beliefs were those that drove her epistemic practices: only members of the Church could be trusted because they were chosen by God, and outsiders could never be trusted. These beliefs would be particularly pernicious because they would have a systematic effect on any belief she would form subsequently. Effectively, any new belief would need to be sanctioned by Church authorities. These new beliefs would, in turn, often have the effect of reinforcing her beliefs about who could and could not be trusted. Despite constantly debating her views with outsiders, Megan’s individual beliefs could survive outside challenges because they were held firmly in place by the rational support of a network of biased beliefs about the world, her religion, and who could be trusted.

She could not correct the influence of the echo chamber one belief at a time. After all, she could only rationally assess any particular belief in light of her total evidence: the intimate web beliefs she already held. It would be irrational for her to believe she was, in fact, not morally justified in her actions without revising the many beliefs that rationally supported her belief she was in the right.

To correct her ignorance, Megan would need to revise her biased belief system as a whole. She would need to attempt to rebuild her beliefs while preventing the beliefs she already held from having any influence on the process. She would need to temporarily suspend all her
beliefs to begin anew, much like Descartes famously attempted in his *Meditations*. Most importantly, she would need to prevent her beliefs about the trustworthiness of any source, or the reliability of any special method for interpreting evidence, from having an effect on rebuilding. She would need to try to trust things as they seem with fresh eyes, and could trust in the testimony of others as long as she was able to treat all sources of evidence with equal default, defeasible trust (see Nguyen, 2018, pp. 16–18). For our present purposes, I’ll assume that it is psychologically possible for an agent to perform such a Cartesian “reboot” or, at least, agents are able to do it well enough to undo the skewing effects of forming beliefs in an echo chamber.

From an epistemological perspective, Megan isn’t blameworthy for skirting any epistemic obligation to reboot her whole belief system. Agents do not have epistemic obligations to perform acts or omissions whenever it is possible to improve their epistemic situation. They can only be obligated to improve their epistemic situation when it is reasonable to expect them to do so (W. P. Alston, 1988; Millar, 2019; Tidman, 1996). Unless an agent has strong reasons to suspect their entire belief system is biased, it is not reasonable for them to attempt the arduous task of revising it in whole. Megan’s belief system, and the epistemic practices they guided, would effectively insulate her from taking any outsider’s challenge seriously. Thus, she would not acquire the relevant reasons merely in virtue of having outsiders tell her her belief system was corrupt. She would have to realize this corruption on her own terms. Eventually, Megan’s faith would begin to crumble when a small group of men forcefully seized leadership of the Church from within. She began to see that the principles that guided her action could be used to justify action she considered immoral or unfair. They were being used to cause harm to her and other female members of her family. Unable to reconcile her faith with her doubt, and fearing for her safety, she ran away from home. Shortly after her escape, she published a letter online asking
those who had been hurt by the Church for forgiveness. Yet, it was clear she had not yet realized
the extent of her distorted worldview. The social media butterfly would stay out of the public eye
for two years as she carefully reassessed her beliefs.

While it may have been possible for Megan to perform a Cartesian reboot to correct her
ignorance prior to her doubts, she did not have much reason to believe doing so would improve
her epistemic success, and she faced no epistemic obligation to attempt to revise her system of
belief. At this point, critics may object that Megan faced a moral obligation to correct her moral
ignorance (e.g., Harman, 2011), and point out that moral obligations are not so easily dismissed
due to practical demands. We may be always obligated to improve our moral situation regardless
of the difficulty. But the barriers Megan faced in overcoming her ignorance were not simply
practical. Megan was not morally responsible for failing to correct her moral ignorance prior to
having her first doubts because she did not have the capacity to do so rationally.

Neil Levy (2009, 2016) has argued that an agent is morally responsible for their ignorance only if (a) they could reasonably have been expected to avoid or correct their ignorance, (b) they were capable of correcting it, and (c) they were afforded opportunities to do so (Levy, 2009, p. 735). Levy argues that our normative demands would be out of place if the correction required some chance event (like a hostile takeover) or “glitch” in their reasoning. He explains, it is reasonable to hold an agent accountable to a normative standard only when they are capable of rationally conforming their behavior to that standard. The agent who is unable to recognize an opportunity to improve their epistemic situation fails to have the rational capacity to correct their ignorance and, thus, isn’t blameworthy for not acting on that opportunity (2016, p. 264). If Megan had clear reasons to revise the system of beliefs that made up her worldview, and failed to attempt to correct her worldview, she would be blameworthy. But, the offending system
of beliefs (and the practices they supported) would prevent her from having any reason to think her epistemic situation could be improved. She didn’t shy from questioning her beliefs because it might result in some inconvenient moral truths; in fact, she was dogged in her attempts to figure out moral truths. Ultimately, she was unable to appreciate that she could not rationally evaluate her beliefs in an unbiased way: that she needed to throw away all her available evidence and start again. Once she started to doubt her ability to evaluate her beliefs and choices, she did not avoid the discomfort. Nonetheless, her belief system would continue to operate defensively. In the months before Megan left the Church, she worried these doubts might be part of a test from God. The fear of eternal damnation might keep most in line, but Megan’s moral concern was stronger than her fear. Her unwavering concern for morality drives her today. She has dedicated her life to combating hate speech, educating the public about extremist groups, and helping to deradicalize others who formed their worldviews in the confines of an echo chamber.

Social practices can render us morally ignorant, and they can render us unable to rationally correct that ignorance. An agent isn’t fully responsible for their false moral beliefs if they are unable to rationally correct them. I’ve argued Megan Phelps-Roper isn’t fully responsible for her ignorance because the epistemic practices of the Westboro Baptist Church prevented her from rationally correcting her moral ignorance. If ignorance can excuse wrongdoing, we should not think she’s fully responsible for the harm she caused others when she believed she was morally justified in causing that harm. I now turn to consider objections to this account.
2.3 - Objections

The primary aim of this chapter had been to argue that social practices can render agents morally ignorant and that agents are not always responsible for this ignorance. I’ve shown how the social practices of the Westboro Baptist Church created conditions where Megan Phelps-Roper was morally ignorant. She falsely believed she was justified in her wrongdoing, yet she was unable to correct her moral ignorance without also correcting a large number of beliefs that rationally supported that ignorance. It may be true that Megan had some moral obligation to correct her ignorance, and that her conduct is the fitting object of a variety of moral appraisals; however, Megan was not responsible for any failure to correct her moral ignorance. She was not capable of rationally conforming her behavior with such a moral obligation as long as her belief system biased her ability to rationally evaluate her beliefs and, ultimately, prevented her from gaining reasons to suspect her belief system as a whole was biased.

Megan’s worldview was rationally insulated from outside correction by the social practices of the Westboro Baptist Church. Although the Church was an echo chamber, members are not typically physically isolated from the information sources that could help them correct their ignorance. Strikingly, their isolation is psychological and rational. I’ll close this chapter by considering how Megan’s situation could illuminate discussions about culture and ignorance on a broader scale. But, presently, I’ll attend to objections that might be made to the account presented.

2.3.1 - But, Isn’t Moral Knowledge Easy?

Michelle Moody-Adams argues against the *inability thesis*: the claim that social practices can render agents unable to know right from wrong (1994, p. 293). To motivate this idea, she
considers whether Ancient Greek slave owners might be excused for owning slaves if the practice was a widely accepted part of their culture. This case might encourage readers to share her intuition that the Ancient Greeks were uncritically committed to preserving their way of life. Many of us may struggle to imagine being in circumstances where we are rendered insensitive to the wrongs of chattel slavery. Wasn’t the evidence of their wrongdoing clear on the faces and bodies of their slaves?

Part of what makes Moody-Adams’ analysis of Ancient Greek slavery attractive is the lack of any explanation of structural ignorance. As she presents the case, there are few reasons for us to think slaveowners were morally ignorant and this ignorance is explained by social practices. The only practice mentioned is that of owning slaves. Presumably, the details of the social practices said to explain an agent’s structural ignorance matter when we’re attempting to determine whether an agent could have done anything to address their ignorance. A plausible explanation for their ignorance might appeal to how they conceptualized or justified slavery. But, if the Ancient Greeks were structurally ignorant, recreating this would be quite difficult. The historical record on Ancient slavery is notoriously disjoined and fragmented (R. Alston, 2011; Finley & Shaw, 1980). How Ancient Greek societies conceptualized the practice of slavery is a matter of ongoing historical debate.

If Moody-Adams’ case is missing key details, it may be because her argument doesn’t require us to attend to them. Her claim that no social force can render an agent unable to know right from wrong seems to be motivated by the idea that moral knowledge is always easy to acquire. She states,
“...in virtue of learning a language, every human being has the capacity to imagine (to conceive) that her social world might be organized on quite different principles... one has the capacity to question existing social practices merely by virtue of learning to form the negation of any statement.” (Moody-Adams 1994, 296 n14)

Moody-Adams seems to claim that we could gain all we need to effectively question the existing social order through the exercise of our imagination. Because social forces can’t damage this capacity, they can’t render us unable to know right from wrong. But there are a few steps missing in Moody-Adams’ moral epistemology.

First, our ability to effectively question the moral worth of a social practice doesn’t reduce to our ability to negate any value statement it implies. Megan Phelps-Roper could imagine a world in which her attempts to convert sinners was somehow wrong, and she could imagine a world in which adultery and homosexuality were morally permissible. But these imaginings would have offended her moral sensibilities. These imaginings might only give her reason to redouble her commitment to her corrupt moral outlook. Social practices may not have corrupted Megan’s ability to perform logical operations, but they certainly affected how she evaluated various moral propositions and states of affairs.

Second, an agent who is capable of questioning a social practice is not automatically able to ascertain the moral worth of that practice. Megan Phelps-Roper was certainly capable of questioning the practices of the Westboro Baptist Church, and did so. But because Megan’s only tools for inquiry came from the Church, her inquiry always resolved in favor of the Church. The process of coming to know typically involves some scope of investigation -- asking the right
questions, gathering and weighing evidence, and figuring out what to believe. At each step, Megan’s commitment to the social practices of the Church biased her inquiry.

The fact that social practices can bias our moral sensibilities or our ability to inquire might not matter to Moody-Adams if she subscribed to some form of moral intuitionism. An intuitionist might argue that an agent only needs to reflect on a moral proposition to know whether it is true or false -- they need no additional evidence or argument. Thus, if a structurally ignorant agent could imagine their social practices were arranged according to some other moral values they would be forced to consider the moral facts. They would be unable to avoid confrontation with the truth.

Intuitionism about moral knowledge is defensible when it comes to very basic moral generalizations. Moody-Adams might be able to maintain that social practices could not render agents ignorant of basic facts like “harm is wrong”. But few intuitionists would argue that all moral knowledge can be acquired through reflection. The intuitionist stance is not easily defended when it comes to more specific forms of moral knowledge. Making judgments about the moral value of a practice may require considerable empirical knowledge. For example, we might want to know whether the conditions in overseas factories or prisons count as slavery, or whether our dependence on them is morally permissible. Our ability to come to a conclusion about the moral permissibility of these situations may be purposefully obstructed by entities that have a vested interest in our not knowing. Our ability might also be hampered by how we conceptualize persons in the “third world” or “criminals”. As long as some moral knowledge requires empirical knowledge, social practices can obstruct an agent’s ability to have moral knowledge.
Megan appeared to have had sufficient access to the basic moral principles to understand it was wrong to cause harm to the grieving families at the fallen soldiers’ funerals. She suspected the Church’s practice of insulting them was wrong, so she asked those she trusted if their practices were morally permissible. They offered her an argument based on ideas and principles she already accepted. Megan’s false belief in the moral permissibility of protesting funerals were justified, in part, by her false belief that she was helping them avoid the eternal pain of damnation. In turn, this belief was based on others concerning scriptural interpretation and who to trust about matters of God’s will. These are among the beliefs that hampered Megan’s ability to correct her ignorance. Megan could not simply intuit her way to the right answers here.

If structurally ignorant agents are to blame for their ignorance, it’s not because gaining moral knowledge is easy. Social practices may not interfere with agents’ abilities to know very basic moral facts, but they can interfere with agents’ abilities to evaluate the moral status of their actions. In Megan’s case, the practices she learned regarding who to trust and how to manage evidence skewed her inquiry to such an extent that rationally correcting her ignorance required revising her belief system as a whole.

2.3.2 - Aren’t They Just Willfully Ignorant?

Though Moody-Adams’ arguments from moral epistemology fall flat, her argument that structurally ignorant agents are merely willfully ignorant is more compelling. This is especially true when the structural ignorance under consideration is entwined with systems of oppression, like racism or sexism. Those writing in the epistemic injustice literature frequently claim that agents enjoying social privilege suffer from willful ignorance, regardless of whether their false
beliefs are part of broader ideologies of thought (e.g., Linda Martin Alcoff, 2007; Mason & Wilson, 2017; Pohlhaus, 2012).

Moody-Adams argues that whether an ignorant agent is rendered ignorant by their culture, or is willfully ignorant, is not a matter that is simply resolved by empirical investigation. The issue here is larger than the fragmented record we have on Ancient slave owners. She argues that the truth, or explanatory value, of any psychological theory cannot be determined on the basis of empirical evidence alone (Moody-Adams, 1994, p. 294). This is true for any theory in the natural sciences (see Duhem, 1991; Quine, 1975, 1976). Philosophers of science disagree about the epistemological ramifications of this problem of underdetermination (Turnbull, 2018); nonetheless, they broadly agree that, when two theories that are equally supported by the empirical data, we must give weight to nonempirical epistemic considerations when choosing which to believe. Some philosophers have suggested we evaluate theories according to virtues they might have (e.g., Lipton, 2003). For example, we may find greater epistemic success by choosing theories that are simpler or that better cohere with other theories we believe. Moody-Adams argues psychological theories ought to

“...be capable of conformity with self-conceptions capable of withstanding rational scrutiny… [and] should be equally attentive to the tendency of some psychological theories to produce self-conceptions that influence desires, ends, and action in a dangerously self-deceptive way.” (Moody-Adams, 1994, p. 297)

By “self-conception” Moody-Adams means our beliefs about ourselves, our place in the natural world, and our relationship to other people. Moody-Adams explains that our self-
conception should recognize ordinary agents’ potential for serious wrongdoing and our self-deceptive tendencies (Moody-Adams, 1994, pp. 298–300). Such a self-conception might favor the hypothesis that, for any given case of moral ignorance and wrongdoing, an agent is more likely to be willfully ignorant than rendered ignorant by social practices. The fact that a theory coheres with what we already know is a clear epistemic virtue: it is rational to believe in theories that cohere with all the available evidence. But Moody-Adams can’t claim a clean win here. The hypothesis she defends is not merely that willful ignorance commonly accompanies wrongdoing, it’s the strong hypothesis social practices do not render agents morally ignorant. Provided what we know about Megan’s case, and other cases where agents can rationally maintain radical beliefs (Hardin, 2002; Nguyen, 2018; Sunstein & Vermeule, 2009), Moody-Adams’ strong hypothesis appears to do a poor job of cohering with what we already know about ourselves.

The second virtue Moody-Adams cites for theory choice may seem more promising. If we accept that agents can lack responsibility for their wrongdoing because they are “victims” of their culture, we may make the mistake of failing to hold blameworthy agents responsible. Ultimately, we may fail to have the right degree of moral concern if we allow ourselves to think we are blameless for engaging in popular social practices. For example, we might be less likely to consider the moral ramifications of our consumer behavior if we can rationalize our behavior by appeal to widespread social norms. This all could be true, yet it would fail to have any bearing on the epistemic merit of a psychological theory. Like many purported virtues of good theories (Barnes, 1995), the fact a theory might have a dangerous influence on the public doesn’t reduce the explanatory power of that theory.
Moody-Adams’ arguments from theory choice fail to show that the inability thesis is false. Social practices can render agents rationally unable to correct their ignorance and, due to this inability, it is unreasonable for us to hold them accountable for not correcting that ignorance. This is hardly a matter of structurally ignorant agents choosing not to be informed about what they can and should know (paraphrasing Moody-Adams, 1994, p. 301). But perhaps Moody-Adams’ notion of affected ignorance is more sophisticated than our ordinary conception of willful ignorance.

Moody-Adams employs a variety of cases to illustrate her accusation against structurally ignorant agents: the torturer that uses tactical language to mask the nature of her violent ways; the head of a bank who insists on not knowing the methods by which his subordinates make impossible profits; the mother who doesn’t ask how her teenage son is able to afford to give her expensive gifts; and the school administrator that refuses to investigate allegations their colleague is guilty of harassment (paraphrasing Moody-Adams, 1994, pp. 301–2). Jan Wieland (2017) uses these very cases to develop a conceptual analysis of willful ignorance. In each case, he argues, the agent claims ignorance of some proposition, p, that implies that their action is wrong. While the agent should and could have considered p, they did not consider p because doing so is “inconvenient”. This inconvenience is either a matter of self-interest -- e.g., motive to preserve their self-image, or to continue immoral behavior that benefits them -- or as a matter of the interest they have invested in others. (Wieland, 2017, p. 7).

Given Wieland’s account, we can distinguish cases like Megan Phelps-Roper’s from genuine cases of willful ignorance. Megan should and could consider her actions were wrong, and she considered the moral status of her actions as best she could. Willfully ignorant agents might appear to consider the moral status of their actions but, having some suspicion they are
engaged in wrongdoing, they avoid a sincere consideration they may be wrong. Key to
distinguishing willful ignorance from ignorance caused by other factors is that the willfully
ignorant agent’s avoidance behaviors are caused by convenience motives. It was inconvenient
for Megan to question the practices of the Church: if she learned she was wrong, she would have
to cope with the fact that she and everyone she loved were routinely engaged in wrongdoing.
Considering that core among Church beliefs was the idea that members enjoyed special moral
and religious status, these convenience motives were very strong. But Megan felt a duty to
question Church practices and never shied from the opportunity to do so. She remained ignorant
because she did not have the capacity to evaluate the moral (or epistemic) status of her actions in
an unbiased way. Her avoidance behaviors were not caused by motives to maintain her self-
image but because they were what she believed was morally and rationally demanded of her.

We can compare Megan to another Westboro Baptist Church member who was clearly
willfully ignorant: Steve Drain. Steve joined the Church at the age of 35 after he had decided to
film a documentary exposing the Church’s wrongdoings. He and his family are the only
members to join the Church that don’t have ties to the Phelps family by blood or marriage.
Steve’s daughter reports he was a staunch atheist before he discovered the Church, and that he
found Church membership attractive because it gave him an opportunity to exercise his need for
dominance and anger -- a need previously expressed in domestic violence. Steve seized
leadership of the Church when Fred Phelps’ health was in sharp decline, and led the take over
that precipitated in Megan’s departure from the Church. (Dickson, 2014; Drain, 2013). Both
Steve and Megan seemed equally sincere in believing they had moral justification to engage in
wrongful Church practices. Because of this, there’s some ground to claim that Steve is genuinely
morally ignorant. But he’s clearly responsible for his ignorance: he intentionally adopted the
beliefs and the practices of the Church as a practical means to gain power. Moreover, Steve took on false moral beliefs and relied on the testimony of Church leaders while he had sufficient reason to reject it all. He had formed a great number of beliefs outside the Church, and he believed the Church’s wrongdoing was significant enough to film a documentary about them. Steve’s outsider status meant he had the epistemic resources to avoid and correct false moral beliefs and other biasing beliefs -- resources that were unavailable to Megan.

Steve Drain is responsible for a variety of benighting acts, the resulting moral ignorance, and freely acting on this ignorance. If he takes any opportunity to question the moral status of his actions it may only be to save face: to retain his power in the Church, his public image, and his belief in his moral superiority. But understanding Steve’s willful ignorance is not a simple matter of observing his current behavior. When defending his actions or his beliefs, Steve’s observable behavior is nearly identical to Megan’s when she was a member. The key difference is in how he came to acquire the false moral beliefs and his capacity to rationally correct his ignorance. The difference between willful ignorance and a case of structural ignorance like Megan’s may not be easily resolved by simple empirical observation; nonetheless, we would fail to understand just how complicated Megan’s epistemic situation is if we simply judged she was willfully ignorant. In fact, it appears Megan was only able to begin to come to grips with the extent of her corrupted outlook with the help of outsiders who were committed to believing that her wrongdoing wasn’t simply a matter of ill will. Insofar as similar states of structural ignorance may explain other important cases of moral ignorance, we may need to fight the temptation to diagnose ignorant agents as willfully ignorant if we want to help them achieve moral understanding.

It may be true that willful ignorance commonly accompanies wrongdoing, and that agents are more likely than not to be willfully ignorant and able to correct their ignorance. It may also
be true that recognizing that structural ignorance can occur poses some kind of moral hazard -- allowing us to deceive ourselves about what we can and should do. Nonetheless, social practices can render agents morally ignorant and unable to rationally correct this ignorance.

I’ve argued that social practices can interfere with agents’ abilities to evaluate the moral status of their actions and in ways that diminish their moral responsibility. I turn to one last objection to my account: Megan’s inability to rationally correct her ignorance doesn’t imply she was unable to have moral knowledge. Because it was possible for her to know, it could be argued she’s blameworthy for acting on her ignorance.

### 2.3.3 - But Aren’t They Blameworthy Anyway?

Despite the fact that Megan might not have been able to rationally correct her belief, some may maintain she’s blameworthy. After all, she was able to come to the right moral beliefs eventually. It might be argued that Megan’s inability to rationally believe the moral truth doesn’t excuse her wrongdoing. Elizabeth Harman (2011) defends such a view, extending Moody-Adams’ arguments to argue moral ignorance is never excused. But Harman gives a more explicit explanation for why the morally ignorant are blameworthy -- appealing to an account developed by Nomy Arpaly (Harman, 2011, p. 460). Arpaly argues a person is blameworthy for being inadequately responsive to moral reasons (2002, pp. 79–80). Arpaly explains,

I take a person to be responsive to moral reasons to the extent that she wants noninstrumentally to take courses of action that have those features that are (whether or not she describes them this way) right-making and not to take courses of action that have
those features that are (whether or not she describes them this way) wrong-making features. (Arpaly, 2002, p. 79)

Arpaly goes on to explain that an agent may not be culpable for failing some moral duty to make sure they have the right factual beliefs:

...as we do not even know which of our beliefs need checking more than others, and as it is in the nature of some kinds of irrationality that one cannot detect it in oneself, it makes no sense to say that we always have such a duty. (Arpaly, 2002, pp. 109–110)

From these decontextualized quotes, we might think that Arpaly agrees that Megan isn’t to blame for her moral ignorance. It does seem that Megan was responsive to moral reasons in the way described: Megan was intensely concerned about making the right moral choices. The ignorance excuse may be available to Megan insofar as her moral ignorance is explained by various false non-moral beliefs -- about who to trust, how to interpret scripture, and so on. But Arpaly seems to allow agents to have ignorance excuses for their wrongdoing only in the event they are confused about factual matters (2002, 102). Arpaly does not seem to allow agents ignorance excuses when it comes to moral ignorance. She argues agents are blameworthy for wrongdoing that is the product of moral ignorance, and praiseworthy for doing the right thing when they believe they are engaged in wrongdoing. For Arpaly, agents are blameworthy for failing to respond to morally relevant considerations (2002, 80). The agent who believes they are morally justified in their wrongdoing is still wrong, and blameworthy for failing to respond to the right moral reasons (2002, 99). Arpaly suggests that agents who falsely believe they are doing
the right thing suffer from ill will: they are either acting on sinister reasons or are indifferent to moral reasons (2002, 101). Matthew Talbert shares Arpaly’s view, and elaborates,

> [e]ven if a wrongdoer is ignorant of the fact that her behavior is wrong, and even if this ignorance is not her fault, her actions may still express the contemptuous judgment that certain others do not merit consideration, that their interests do not matter, and that their objections can be overlooked. (Talbert, 2013, p. 324)

While Arpaly doesn’t explicitly argue every case of moral ignorance involves ill will, Harman (2011) does. But Harman argues this is because agents are always in a position to realize the moral facts (Harman, 2011, p. 463). I’ve argued agents are not always in a position to gain moral knowledge (§2.3.1). Not all moral truths are available by mere reflection. Even so, I’ve argued social practices can interfere with an agent’s ability to gain truth through reflection. Megan was rationally unable to determine that she had false moral beliefs. Further, I’ve argued social practices can interfere with an agent’s general ability to inquire. Megan’s inquiry was biased by background beliefs that made it irrational for her to gain true moral beliefs. But, perhaps, agents are always in a position to realize the wrong-making features of their actions. This doesn’t require belief. As promising as this is, it’s patently false that agents always have access to the wrong-making features of their actions. The moral consequences of many of our choices are not immediately available to us. This is especially true when it comes to our consumer behaviors, where it may be impossible to find out whether an “ethically sourced” product is actually ethically sourced (e.g., Emerson, 2019).
Agents may not always be in a position to gain moral knowledge in a way that supports Harman’s generalization; nonetheless, Megan had access to the wrong-making features of her actions. She recognized she was causing others harm regardless of her beliefs. It’s unclear how we should understand moral justification on an account like Arpaly’s or Harman’s. Presumably, in genuine cases where an agent has moral justification the agent acts on right-making features that make them blameless for also acting on some wrong-making features. The good would, in some way, balance out the bad. Since Megan wasn’t motivated by genuine right-making features, only a false belief, perhaps her actions are just as blameworthy as they would be if she simply ignored the pain she could see in the funeral-goers faces. This might make her ill willed on Harman’s account.

When an agent appears to be indifferent to moral reasons, or seems to have sinister reasons for acting, it seems sensible to say they are ill willed. The fact they believed they had moral justification wouldn’t compel us to think they actually cared about morality. Perhaps a great many people suffer from ill will, in which case, it doesn’t matter much whether they are structurally ignorant. They wouldn’t bother to gain moral clarity even if they had the opportunity. But, once Megan was in a position to rationally correct her structural ignorance, it seems her moral concern compelled her to gain clarity. If Megan is nonetheless blameworthy for “ill will”, the claim doesn’t appear to amount to anything more than the fact she is guilty for being inadequately sensitive to the wrong-making features of her actions. Using the concept of responsibility I’ve operated with thus far, it still makes sense to ask if she’s responsible for failing to be moved in the right way. As long as this sensitivity was shaped by her false moral beliefs, we should maintain she isn’t responsible for her moral ignorance. If Megan is blameworthy, this judgment is distinct from my claim that she isn’t responsible.
Arpaly’s view of blameworthiness appears to be informed by the account of blame developed by P. F. Strawson (1974). Strawson criticized accounts of responsibility -- like the one developed here -- for failing to account for our actual practices of holding people responsible. Specifically, it neglects the significance of the affective responses we have to interpersonal injury and how these responses are part of the framework of ordinary moral life. Strawson argues blame is a reactive attitude: “natural human reactions to the good or ill-will or indifference of others as displayed in their attitudes and actions” (Strawson, 1974, p. 53). Negative reactive attitudes -- like blame and resentment -- express a demand for moral consideration. In order to simply evaluate whether Megan Phelps-Roper is blameworthy despite my arguments so far, I’ll bracket traditional concerns about whether Strawsonian accounts that describe our practices of blame are suitable accounts of how we should blame (cf. G. Watson, 1987, pp. 144–8). If Megan is blameworthy on this account, Harman and Moody-Adams might be able to maintain the claim that structurally ignorant agents are always blameworthy for their ignorance.

Megan might be blameworthy for her demonstrated failure to give adequate moral consideration to others’ interests, regardless of her beliefs. But, Gary Watson argues, the reactive attitudes aren’t merely reactions to an agent’s quality of will: these attitudes respond to our background beliefs about the agent being judged (G. Watson, 1987, pp. 124–6). Because our blame involves a moral demand, our attitudes change when we realize an agent only appeared to fail to meet that demand. If our attitudes are inhibited by such an explanation of their behavior, the agent has a valid excuse or was justified in their action (G. Watson, 1987, p. 123). Excuses deny the appearance the agent failed to meet the demand for moral regard. Thus, we might cease to feel it appropriate to blame our friend for breaking their promise to drive us to the airport when we find out they had to rush their child to the hospital. They failed to keep their promise,
but they aren’t to blame for displaying ill will in their (lack of) action. Some cases of factual ignorance might inhibit our blame on the Strawsonian account insofar as an agent’s failure to know isn’t a display of ill will or indifference. These may be the most common kinds of ways agents defend against blame in ordinary moral life: “I’m so sorry. I didn’t know!” But, considering Harman and Moody-Adams seem to think we commonly engage in self-deception, they might advise us to be far less forgiving about factual errors.

Moral ignorance seems always to involve some failure to meet interpersonal moral demands for moral consideration. There is no way of denying the appearance the ignorant failed to meet these demands: no excuse. Nonetheless, certain kinds of moral ignorance do mitigate our tendency to blame. Blame seems to only make sense insofar as the agent can, in some way, be held accountable to the demand. While the sensation of stepping on LEGO might automatically provoke the urge to blame the toddler who left it there, there’s something off about blaming the toddler. In terms of a communicative act, the demand for moral regard communicated in blame is lost on the tiny human who is just learning how to communicate. In terms of the demand implied in the attitude, blame fails to make sense when the agent subject to it doesn’t yet have the moral understanding necessary to be responsive to the demand (G. Watson, 1987, p. 126). Agents who can’t be held accountable in this way are exempt from the moral demand. We may be able to justify having certain reactive attitudes towards toddlers insofar as it might better shape their moral learning (e.g., Fricker, 2010; McGeer, 2019). Nonetheless, we would make a mistake to believe we were using blame on the toddler in the same way we would use blame with adults. When we blame the adult, we are holding them to account for their actions. With the toddler, we’re treating them as something that needs to be controlled -- not someone who can be reasoned with.
Watson considers how Strawsonian accounts of blame struggle to make sense of how we engage with grown adults who seem to be exempt. If another agent rejected or flouted our demand, they seem to have moral understanding (G. Watson, 1987, p. 130). But an agent’s moral sensibilities may be so thoroughly corrupted that our demands are unintelligible to them. Strawson claims we are more inclined to take an *objective attitude* when we evaluate agents who are deranged, extremely compulsive, or are otherwise the product of an unfortunate upbringing. This objective attitude extracts us from interpersonal involvement with the other agent. We no longer treat them as agents responsive to moral reasons, but as the mere products of unfortunate causes; not as someone we can reason with but, like an animal, something to be controlled.

Using a Strawsonian account of blame, it seems we should judge that Megan Phelps-Roper was exempt from our blame. She demonstrated she was not capable of being responsive to interpersonal moral demands, though she was confronted with outsiders’ negative reactive attitudes towards her daily. Though she did eventually correct her ignorance, it was not by any process of reasoning: she was hurt and confused, and fled the Church. Unable to reconcile her doubt and her faith, she was forced to question everything she thought she knew. The fact that Megan appeared to be so morally concerned while with the Church might seem to complicate this result. But this is only because we are confusing her concern to have moral justification for genuine concern over the wrong-making features of her actions. Arpaly’s account of blameworthiness gives us a different way of looking at Megan’s moral psychology; nonetheless, Harman and Moody-Adams appear to simply deny the possibility that an agent may be exempt from any moral demand to consider the moral interests of others. They deny this when they claim that agents are always in a position to gain moral knowledge. Both fail to account for the way social forces can render agents rationally unable to correct for their moral ignorance.
If the moral understanding Megan has today means she is no longer exempt from blame, it doesn’t imply she’s to blame for her past behaviors. Megan acknowledges there are people who still resent her for the harm she caused them in the past. Blaming Megan for what she did in the past fails as an expression of a demand for reasonable regard insofar as she already repudiates her past conduct (cf. G. Watson, 1987, p. 129). The blame expressed might make sense as a form of an address to a Megan who no longer exists, but the demand wouldn’t have been intelligible to her anyway. Nonetheless, Megan appears to be taking responsibility for her past actions insofar as she seeks forgiveness. But agents can take responsibility for another’s actions without being morally responsible for those actions. A parent may take responsibility for the harm caused by their child without necessarily being to blame for their child’s outburst. Thus, there may be room for some people she offended to hold negative reactive attitudes towards Megan insofar as these attitudes communicate a demand for reparations. These attitudes can be consistent with the recognition that she wasn’t fully responsible.

Watson notes that we may experience conflicting attitudes with cases like Megan’s where an agent’s upbringing explains their morally corrupt outlook (G. Watson, 1987, p. 138). We may understand her as a “victim” of her circumstances, and struggle to believe we could have done any better in the same situation; nonetheless, this objective attitude may not dispel our outrage towards her role as a victimizer. We may cling on to our reactive attitudes, despite the fact that the demands they make seem pointless and even though they may do more to damage our interpersonal relationships. Ultimately, Watson shows how tensions like these make it difficult to defend the idea that the Strawsonian analysis of how we place blame should guide us to how we ought to blame (G. Watson, 1987, pp. 138–148). Watson maintains, “[t]he boundaries of moral responsibility are the boundaries of intelligible moral address” (G. Watson, 1987, p. 148). If
social practices interfere with an agent’s ability to make the moral demands on them unintelligible and keep the agent from rationally correcting their ignorance, we could not reasonably demand (or expect) they know or do better.

2.4 - Conclusion: Beyond Brainwashing

I argued that social practices can render agents morally ignorant (§2.2). Systems of belief shared in communities can rationally justify false moral beliefs. But worse, social practices of asking questions, trusting, and interpreting information can render agents unable to rationally correct their ignorance. Correcting their ignorance requires they correct for the biasing influence of a great number of their beliefs. It’s unreasonable to expect structurally ignorant agents like Megan to gain moral knowledge when it’s irrational for her to do so. When an agent can’t rationally conform their behavior to some normative standard, they can’t be responsible for their failure to meet it.

I then considered three objections to this view. I reject Moody-Adams’ arguments that agents are responsible for their moral ignorance because agents always have access to moral knowledge (§2.3.1). I also reject Moody-Adams’ argument that we should believe morally ignorant agents are always willfully ignorant (§2.3.2). Finally, I consider the view that agents might be blameworthy despite their inability to rationally correct their corrupted moral outlook (§2.3.3). I concluded that, even if structural ignorance fails to excuse an agent’s moral ignorance, their inability to be adequately responsive to moral reasons makes them exempt from blame.

Megan’s case is certainly special: most people don’t grow up in cult-like situations. Nonetheless, the warped epistemology of echo chambers may be regularly found in contexts of group conflict. If you grow up in a predominantly white suburb of a large U.S. city, everything you know about black Americans might come from white people. Practices that morally justify
discriminatory behavior (Bonilla-Silva, 2006; C. Mills, 2007) may be reinforced by practices of 
distrusting and discounting the very people that might correct your misunderstanding (Dotson, 
2014). These echo chambers are further reinforced by physical practices of social segregation. 
Moreover, when prevailing prejudices are reinforced in popular media, and in the accumulated 
consequences of a long history of racial disparity (Cambria et al., 2018; Mallon, 2016, Chapter 
6), it might not be surprising if false beliefs about race and racial justice are rationally resistant to 
correction. What such a situation might mean for responsibility, and for ameliorating conflict, is 
worth deeper investigation.
Chapter 3 - Do Our Social Identities Limit What We Can Know?

3.1 - Introduction

We often find ourselves in disagreement with others and we may often have good reasons not to change our views. But there are occasions where we gain strong reasons to think our beliefs are false, such as when we realize our interlocutor has an *epistemic advantage*: compared to us, their beliefs about the matter are more likely to be true. There are many factors we might use to determine whether our interlocutor has epistemic advantage. For example, they may have more relevant background knowledge on the subject, or they have better evidence for their claims. These carry clear epistemic consequences, thus give us a clear reason to reconsider our position (Frances, 2014, Chapter 4).

There are some social categories that we might use to reliably identify that our interlocutor has an epistemic advantage. If some social category C can be used as a reliable indicator of an epistemic advantage, then

(i) members of C are more likely than not to have property \( p \)

(ii) \( p \) explains why members of C are more likely to have epistemic success in domain D than non-members

Broadly, epistemic successes include justified beliefs, knowledge, and understanding. When it comes to diagnosing joint pain, orthopedists have certain clear epistemic advantages because (i) their membership in the group is standardly conferred by a process of training and certification and (ii) this process explains why they are likely to be correct in their diagnoses of joint pain. There may be orthopedists that maintain their title through a matter of luck or
deception; nonetheless, the title is a reliable indication they have the relevant certification and training. The fact they are an orthopedist doesn’t guarantee that every diagnosis they make is correct, but it does mean they are more likely than non-members to be correct. Thus, most people would be reasonable to trust an orthopedist when it comes to caring for their joint pain. A number of other healthcare professionals that may also have a clear advantage when it comes to judgments about joint pain. Not all healthcare professionals can be considered experts on joint pain, but their training may still give them an advantage over the general public.

What about social categories like gender or race? In political activism, and politically-engaged scholarship, we may encounter a variety of claims that an agent’s social identity provides the basis for epistemic advantage. Some claims have modest epistemological commitments. Perhaps women committed to engaged feminist activism are better at identifying subtle forms of sexist behavior than men who only have a theoretical understanding of sexism (e.g., Jones, 1999 esp. 65-66). Here, the claim about women’s epistemic advantage is modest because (1) the scope of advantage is specific and, (2) if there were a regular difference in the reliability of men’s and women’s judgments, it seems men are capable of gaining more reliable judgments through activist work. In theory, the property that explains women’s advantage is available to men. Modest claims like these contrast with more radical epistemological commitments. Radical claims about women’s epistemic advantage can be found in the early writings of Standpoint theorists. Sandra Harding writes,

When people speak from the opposite sides of power relations, the perspective from the lives of the less powerful can provide a more objective view than the perspective from the lives of the more powerful (S. Harding, 1991, pp. 269–71)
In a similar spirit, Nancy Hartsock writes,

...there are some perspectives on society from which, however well intentioned one may be, the real relations of humans with each other and with the natural world are not visible (Hartsock, 1983a, p. 117)

Both Harding and Hartsock were highly influential philosophers in the development of Standpoint Theory. Other notable figures included the sociologists Dorothy Smith (1974) and Patricia Hill Collins (2002). While there were important disagreements among Standpoint theorists (see S. G. Harding, 2004), they broadly agreed on three core claims. First, social identities -- like gender, race, and ethnicity -- are epistemically significant and, second, oppressed or marginalized groups are afforded epistemic successes that are unavailable to socially privileged groups (in principle or in practice). Third, because of this epistemic difference, certain lines of inquiry ought to start from the lived experiences of marginalized groups. Due to heated controversy over the first two claims, Standpoint theorists have largely abandoned attempting to justify them. Instead, contemporary theorists have turned to developing moral and political grounds to justify the third, methodological claim (S. Harding, 2009; Rolin, 2009; Rouse, 2009).

Standpoint theorists made radical claims about epistemic advantage. The claims were radical on two counts. First, they brought into question large domains of knowledge -- scientific knowledge (Harding) and sociological knowledge (Smith, Hartsock). If their analyses were correct, we would need to revise a considerable amount of what we take ourselves to know and reform institutions of knowledge production (like sociology departments). Second, these claims were radical in claiming there were epistemic “chasms” between groups of people: for example,
that women have access to knowledge that men don’t. This second claim puts men in a peculiar epistemic situation when they disagree with women. When an orthopedist rejects a patient’s self-diagnosis, the patient has access to clear reasons to change their mind. In theory, the patient could go through the relevant medical training and verify that the orthopedist is correct. In practice, patients rely on semi-governmental institutions to only allow people with this training to claim they’re an orthopedist. By contrast, if there’s an epistemic chasm between women and men, men may not have clear reasons to revise their beliefs when a woman disagrees with them. It may be impossible for him to gain whatever property makes her judgments more reliable and, ultimately, he would be unable to independently verify her advantage (Hekman, 1997; see also Longino, 1993). Even if he has moral, political, or practical reasons to trust her, it may not be epistemically rational for him to do so. Consequently, such a chasm could prove to be a significant obstacle to moral or political progress -- especially if women struggle to gain recognition for their epistemic advantage.

In this chapter, I am interested in investigating whether it’s possible to maintain a radical claim about women’s epistemic advantage. Though Standpoint theorists have left radical claims behind, we may encounter claims that look radical in public debates about women’s issues. I argue that it is possible to defend a radical claim about women’s epistemic advantage: it challenges a large domain of judgments and may be unavailable to men. I will not attempt to generalize this analysis to claims about race, ethnicity, or other social identities; nonetheless, it may be useful in evaluating other claims about epistemic advantage.

To gain clarity on how we might develop a radical claim, I’ll first evaluate hypotheses from Standpoint theorists and work influenced by them. I will not provide a comprehensive history of Standpoint theory or the debates surrounding it, as others have already done this (e.g.,
S. G. Harding, 2004). I am only concerned with determining whether a radical claim to women’s epistemic advantage can be defended. I first consider hypotheses that women have some special property that makes their judgments more reliable. These include arguments that attempt to ground advantage in bodily differences or innate differences in cognition (§3.2). I also consider arguments explaining advantage in terms of gendered differences in social experience (§3.3). If any of these approaches promises to ground a claim that women have more reliable judgments, they fail to support a radical claim: they fail to implicate a large domain of judgments, or fail to show that men aren’t just as capable of these reliable judgments. I then turn to consider a different strategy for developing claims about women’s advantage: by showing that men are regularly epistemically disadvantaged (§3.4). I argue appeals to men’s bias or epistemic vice have promise, but deny they can be fully developed into a radical claim. I close with a proposal for a novel approach to developing a radical claim about women’s epistemic advantage. Women may be distinctively capable of having greater epistemic success than men because oppressive social practices render men’s judgments about sexism unreliable, and render it distinctively difficult for men to correct for these unreliable judgments.

3.2 - Biological Advantage

One method for defending the claim that women have an epistemic advantage over men is by locating a property that is only available to women that affords them distinctive epistemic success. Prior to the earliest formulations of Standpoint Theory, some writers claimed that women have epistemic advantages over men in virtue of having different bodies. Sandra Harding considers a number of thinkers who could be interpreted to make claims that gendered biological functions like menstruation, pregnancy, and menopause might provide a basis for gender
differences in knowledge (S. G. Harding, 1986, p. 180). Linda Alcoff advances similar claims in *Visible Identities* when she argues women’s subjectivity and identity is constituted by a distinct gendered “position” (2005, p. 148) which, in turn, is rooted in embodied differences (2005, p. 172). However, Alcoff might be better understood as ultimately making claims about differences in social experience which non-accidentally track bodily differences. I’ll consider socially-grounded claims in the next section.

As far as we use our bodies to perceive, differences in bodies may lead to some differences in epistemic success. But it’s unclear to what extent these differences are epistemologically significant. Thomas Nagel suggests we may be unable to know what it’s like to be a bat since we lack echolocation (Nagel, 1974, p. 438). In contrast, congenitally blind people demonstrate an impressive understanding of the first-person experience of sightedness (Bedny et al., 2019; e.g., Koster-Hale et al., 2014). Even if congenitally blind people may not have the best understanding of what vision feels like, there’s little reason to think they couldn’t generally have as much or more success than sighted people. With the exception of notable philosophers (like British Empiricists), most accounts of perceptual justification would allow blind people to have justified beliefs about appearances.

It’s possible that women have special knowledge in virtue of their biology. It seems sensible to claim that women have an epistemic advantage when it comes to knowing what it is like to be pregnant. At least, only agents with certain biological configurations can have access to the experiences necessary to know what pregnancy is like. The fact that at least 41 men have had access to these experiences (Light et al., 2014) isn’t particularly troubling to this claim about epistemic advantage as long as women are more likely to have these experiences. But, to use an agent’s status as a woman to reliably judge epistemic advantage here, it needs to be the case that
women are more likely than not to have these experiences. It’s difficult to find reliable data to make this assessment, and there are a few complications. More women are choosing to become mothers later in life (Livingston, 2018), which could mean it’s more likely any given woman has not had the experience of pregnancy; however, any woman with a uterus may be far more likely to have reliable judgments about the experience than men who have never had a uterus.

The line of reasoning here may seem tiresome because it would be foolish to assume someone has experience with pregnancy just because they’re a woman. Nonetheless, this line of reasoning is crucial to show a claim about women’s epistemic advantage could be explained by biology despite the fact that not all women share the same bodily configuration, and despite the fact that “woman” is not just a biological distinction (Godman, 2018). Yet, from the perspective of developing a radical claim about women’s epistemic advantage, biological difference is not particularly promising. The scope of the advantages to find here are fairly narrow. Even if there are epistemic chasms around knowing-what-it’s-like, these don’t seem to pose serious moral or political challenges to men. While most men may never know what it’s like to be a woman, they don’t need this knowledge to be morally sensitive to the distinctive needs most women have by virtue of their bodies. When women’s bodies are at the center of political controversy, it might seem like better moral and political outcomes could be secured if men could live in the body of a woman for a while. But, more realistically, the true barriers to progress are prejudices and social practices that enable men to discount or neglect the claims and ideas stated by women. I consider how this social discrediting might explain epistemic difference in §3.4.

A related, biological approach to defending women’s epistemic advantage has been to claim that there are regular, innate differences in the cognitive abilities of men and women. But
this approach isn’t very promising because the claim to cognitive difference can’t be adequately empirically supported. Despite considerable empirical study on gendered differences in cognition, a strong case can’t be made that there are significant cognitive differences between men and women (Fine, 2010; cf. Margaret M McCarthy, 2011). Even if there were strong evidence (for example) that adult women were better at theoretical reasoning tasks than men, the result would not be sufficient to show this difference isn’t the product of socialization (Code, 1981).

At the moment, there’s very little psychological research that links biological difference to psychological difference. Some findings suggest that sex hormones have interesting effects on cognitive function when they are exogenously introduced, but it’s not clear these tell us much about the effects of endogenous hormones and researchers struggle to replicate any interesting findings in larger studies (Celec et al., 2015; Markou et al., 2005; Nadler et al., 2019). While genetic differences (i.e., genes on sex chromosomes) do seem to explain some dimorphic differences in brain chemistry, the only functional consequence appears to be susceptibility to neurological disorders (Sigurdardottir et al., 2020).

It might seem common knowledge that men and women think differently, and we may be able to locate some differences that explain women’s epistemic advantage; however, the current state of empirical research should lead us to believe these differences have social causes. I now turn to attempts to explain women’s epistemic advantage in virtue of social difference.

3.3 - Social Advantage

While we may fail to find promising claims about women’s epistemic advantage in any biological property that affords women epistemic success, we may nonetheless find women are
afforded epistemic success in virtue of benefitting from some socially-caused property. We could imagine a society where women are encouraged to pursue careers in academic research while men, mostly relegated to manual labor, face social pressure to join the workforce as soon as they graduate high school. In this society, gendered differences in education may be the product of arbitrary and unjust social practices; nonetheless, women would have considerable epistemic advantages as a result of their privileged access to higher education. But, in reality, women have historically faced systematic barriers to education. Even so, some theorists locate women’s epistemic advantage in their social disadvantage.

As a graduate student, Dorothy Smith (1974) experienced moments of alienation as she was studying sociological research -- research that had been exclusively conducted by men. She argued that these men’s explanations of sociological problems were not just wrong because they excluded the perspectives of women, but that women’s experience of oppression gave them special insight into sociological problems (D. E. Smith, 1974, pp. 9–10). In this section, I consider two versions of this claim to special insight: that sexism causes suffering that is epistemically beneficial (§3.3.1) and, relatedly, that women’s distinctive experiences give them special insight into the nature of oppression and how to end it (§3.3.1).

In this section, I examine claims that women have more reliable judgments about matters of sexism because of their experiences with sexist oppression. These claims can be found both in early formulations of Standpoint Theory and in the more recent Epistemic Injustice literature. But there have been notable attempts at social explanations of women’s advantage that do not appeal to oppression. These are claims about socialized gender differences in cognition. Nancy Chodorow (1978) and Carol Gilligan (1993) advance claims that women are especially attuned to interpersonal relationships in virtue of their upbringing. If this is true, then perhaps women
have more reliable access to certain moral facts or reasons. Since morality is important and pervasive, this would be a significant epistemic advantage. Yet, the claim fails to support radical projects if these “distinctive” modes of cognition can be learned or trained, resolving any threat of a chasm. Ultimately, there is little to no empirical support for these claims about gender differences in moral cognition (cf. Brabeck, 1983).

3.3.1 - The Epistemic Value of Suffering

In contexts of oppression, social practices systematically harm members of a social group due to their membership in that group; these harms constitute unjust and unequal constraint, and benefit members of some other social group (Cudd, 2006, p. 25). The group that benefits from oppressive practices enjoy social privilege. Some Standpoint theorists advance the claim that the harm and suffering of oppression is, in itself, epistemically beneficial. Allison Jaggar argues,

…[the] intensity and relentlessness of their suffering constantly push oppressed groups toward a realization that something is wrong with the prevailing social order. Their pain provides them with a motivation for finding out what's wrong, for criticizing accepted interpretations of reality and for developing new and less distorted ways of understanding the world. (Jaggar, 1983, p. 370)

It may be true that most women have some experience of the harms of oppression, and that men lack this experience due to their social privilege. But it is not immediately obvious that the experience of suffering alone makes it any more likely for an agent to develop less distorted ways of understanding the world. Jaggar doesn’t explain this claim further; however, Karyn
Freedman (2006) makes the case that a woman’s experience with sexual violence can cause her to gain a better understanding of systematic violence against women.

Freedman argues that a common symptom of post-traumatic stress disorder (PTSD) is epistemically beneficial: survivors of sexual violence lose the (false) belief that they live in a just world (Freedman, 2006, p. 111). Ultimately, Freedman attempts to argue that survivors are more likely to have true beliefs about violence against women -- in particular, more reliable judgments about the frequency of sexual violence (Freedman, 2006, p. 116). However, it’s concerning that Freedman does not engage with clinical research on PTSD.

Clinical research shows it is characteristic for survivors of sexual violence to be burdened with a variety of cognitive distortions that keep them in a state of emotional distress. In particular, it is common for survivors to falsely believe they are responsible for their assault (Abdullah et al., 2011). Therapies with demonstrated effectiveness aim to rid survivors of their distorted beliefs -- about the event, themselves, and other people. The belief that the world is unsafe is thought to be particularly damaging -- amplifying and maintaining the symptoms of post-traumatic stress (Iverson et al., 2015). Survivors take extreme efforts to avoid a great deal of their normal activities due to having a heightened sense of threat, leading to severe social and occupational impairments. Clinicians interpret this behavior as an inaccurate assessment of the likelihood of future traumatic events, directly in conflict with Freedman’s claim (Ehlers & Clark, 2000).

If survivors of sexual violence have some epistemic advantage in virtue of the harm they experienced, most women wouldn’t have the advantage Freedman describes (Breiding et al., 2014). Freedman’s claim about advantage is radical insofar as it seems the advantage is only available to women, but the advantage is too narrow to support radical projects. The advantage is
small in scope, and it would seem men could share this epistemic success via testimonial
exchanges. While Freedman may not have identified a radical claim about women’s epistemic
advantage, we might adapt her arguments to defend a radical claim. If the nature of oppression is
systematic, it may be reasonable to expect that there are commonalities among women’s
experience of oppression (i.e., experience of harm). Provided these harms don’t cause traumatic
stress or suffering that hindered epistemic success, women’s distinctive experiences of sexism
may, in some way, reveal the nature and prevalence of sexist oppression. José Medina appears to
make arguments along these lines.

3.3.2 - Regularities in Experience

Medina argues that oppressed agents have distinctive experiences which can afford them
a better understanding of oppression. He explains this advantage comes out of the epistemically
significant harms of oppression. For example, oppressive practices may routinely cause the
oppressed to feel intellectually inferior and to have constant self-doubt (Medina, 2013, pp. 41–2).
In *Epistemic Injustice*, Miranda Fricker (2007) describes how these harms can occur through
practices of discrediting, misinterpreting, and silencing (see also Dotson, 2011). Other epistemic
harms are characterized by the systematic exclusion of the interests, values, and perspectives of
oppressed groups from public projects of knowledge production (Collins, 2002; Dotson, 2014).
While the described harms would seem to hamper agents’ epistemic success, Medina suggests
that these harms may be the basis for the development of epistemic virtues -- cognitive and
affective dispositions that make them more epistemically successful. He explains that traits like
intellectual humility, curiosity, diligence, and open-mindedness are the result of demystifying
experiences that can only be had by members of oppressed groups (Medina, 2013, p. 45). The
demystifying experiences he describes occur when members of an oppressed group notice that their desires, feelings, and thoughts are misrepresented (or not represented) by prevailing social practices. He claims these moments form the basis of an epistemic advantage insofar as they enable moments to question received epistemic norms and ways of conceptualizing the world (Medina, 2013, pp. 46–7).

If members of oppressed groups are more likely than not to have these moments of alienation, and they reliably afford epistemic success, then they have a clear claim to epistemic advantage over members of socially privileged groups. The claim is radical insofar as the experience is only available to members of oppressed groups, and insofar as the scope of the advantage is large. Effectively, agents with this advantage have more accurate judgments about the moral and epistemic value of prevailing social practices, and better insight into how to change these practices. Because those with the power and agency to make these changes are “epistemically ill-equipped”, Medina argues they can only fight oppression by working with those with epistemic advantage (Medina, 2016, p. 195).

Medina’s arguments might encourage the following radical claim of epistemic advantage. By virtue of their experience of oppression, women have more reliable judgments about the nature of sexism and the moral/political reform necessary to ameliorate sexist oppression (Hartsock, 1983b; e.g., D. E. Smith, 1974). This may not feel like a particularly controversial claim at first glance, but there is a considerable tradition in feminist thought that rejects this simple claim. As it is often recounted in academic circles, the early feminist thought was led by middle to upper-class white women. Frequently, these women attempted to speak on behalf of all women as they explained the nature of sexist oppression and how best to combat it. But, their claims were hotly contested by black women, who were more frequently from working class
backgrounds (Collins, 2002; Hooks, 2000). These critics denied that white, middle-class women accurately understood sexism at large because they failed to understand how sexism affected black women in different ways. Moreover, they denied early critics could effectively combat sexism without understanding how racist and classist social practices modulated the expression of sexism. These critiques were among the defining features of feminism’s “third wave”. The claim that anyone could have reliable judgments about sexism simply by having some experience of sexism would continue to be challenged by scholars who continued to map the intersections of oppressive systems, (e.g., Mohanty, 1984; Narayan, 2013). Challenging Medina’s suggestion that oppressed groups might reliably develop virtue from their experiences of alienation, third-wave critics noted ways in which women would be driven by their demystifying experiences into epistemic vice: arrogance, laziness, and closed-mindedness (Abu-Lughod, 2002; e.g., Ortega, 2006)

There is substantive disagreement among these third-wave critics of early feminist thought; nonetheless, they shared a common claim: any woman’s claim to insight on matters of sexist oppression are restricted by the limitations of her own experience. Even if members of oppressed groups have distinctive experiences, and members commonly have these experiences, their experiences don’t guarantee them insight into the systemic and systematic nature of oppression. These experiences can only offer them insight into the local effects of a larger phenomenon. Thus, women might achieve a distinctive epistemic advantage when they are able to share their experiences with other women (e.g., MacKinnon, 1989, sec. II). Women may be in a better position to arrive at truths about sexism because their experiences provide distinctive motives for them to engage in collective inquiry in action -- activities that could give them an epistemic advantage. But the fact that this advantage is achieved primarily on the basis of
testimony undercuts any radical claim to an epistemic chasm between men and women. In theory, men have equal access to this advantage.

Though Medina maintains that experiences of oppression can afford a distinctive epistemic advantage, he ultimately denies that we can reliably judge an agent has an epistemic advantage simply because they are oppressed (Medina, 2013, p. 45). But, I’ll argue, there is a good reason to think we can. Oppressed agents aren’t afforded any radical epistemic advantage as a result of having any property that makes it more likely they achieve epistemic success: their advantage derives from the fact that socially privileged groups are at an epistemic disadvantage.

3.4 - Lack of Disadvantage

Attempts to characterize the radical epistemic advantage of oppressed agents most commonly locate the basis for that advantage in a special property held by the oppressed group: distinctive bodies, superior cognition, or enlightening experiences. But advantage is a relative assessment of epistemic success. An agent may have the epistemic advantage in not being epistemically disadvantaged. If women have an epistemic advantage, it may be due to a property they lack: a property that decreases the likelihood of men’s epistemic success. Women’s judgments about sexism may be more reliable than men’s because men are disadvantaged by cognitive biases, epistemic vice, or willful ignorance.

The derivative nature of such a claim may seem to undercut the radical spirit of the claims defended in the Standpoint literature: women don’t have some special power of insight, they are simply less likely to have their inquiries hampered by bias and vice. Despite this departure from the spirit of some Standpoint arguments, I’ll show how we may nonetheless find here a radical claim about women’s epistemic advantage. I consider claims that men’s judgments
about sexism are rendered unreliable by prejudice (§3.4.1) and epistemic vice (§3.4.2). I find
promise for recovering radical claims about epistemic difference. Yet, I argue this promise is
better fulfilled by a more holistic consideration of the way oppressive social practices make
men’s judgments about sexism less reliable, and make it distinctively difficult for men to correct
for their errors.

3.4.1 - Masculine Biases and Prejudices

Men might have an epistemic disadvantage if they suffer from distorted thinking. I’ll
attend to varieties of motivated reasoning later in the section, but it may be the case that men’s
epistemic successes are hampered by cognitive biases. There’s not much research to support the
claim that men are any more susceptible to common biases such as seeking information that
supports what they already believe (confirmation bias)(cf. Traut-Mattausch et al., 2011), using
information most easily recalled when making judgments (availability bias), or at generally
overestimating their abilities (the Dunning-Kruger Effect)(R. M. Watson et al., 2019; cf. West &
Eaton, 2019). But, perhaps, prejudices against women can put men at an epistemic disadvantage.

In *Epistemic Injustice*, Miranda Fricker suggests that assessments of women’s credibility
are systematically undervalued due to identity prejudices (2007, p. 28). She explains that
stereotypes with contents like “women are irrational” (2007, p. 24) negatively influence
judgments of women’s credibility even if the listener believes the stereotype is false (2007, p.
35). She calls this *testimonial injustice*. Potentially, Fricker is concerned with implicit sexist
biases -- automatic cognitive or affective processes that bias thought and behavior in ways that
contribute to the oppression of women (McKinnon, 2016; cf. Saul, 2017). At first glance,
findings from Harvard’s Project Implicit appear to lend some credence to the idea that men are
more likely to be implicitly biased than women across a variety of contexts (Nosek et al., 2002, 2007, pp. 27–8); however, findings don’t support consistent gendered differences in implicit social stereotyping across testing, and women actually have considerably stronger implicit gender stereotyping than men (Nosek et al., 2007, pp. 28–9 & 12). This finding coheres with the hypothesis that oppressed groups are just as likely as privileged groups to have attitudes that preserve the social structures implicated in oppression (Jost & Banaji, 1994). These findings don’t challenge the claim that women are systematically discredited or silenced, but they should discourage us from thinking implicit biases are the source of any gendered epistemic difference.

Men may more likely to suffer from distorted thinking based on the biased beliefs they have about women. These are explicit prejudices. Testimonial injustices caused by explicit prejudices might be quick and unthinking but, if challenged, the listener would maintain they had good reasons not to trust a woman’s testimony. While women and men may be just as likely to share _benevolent_ sexist beliefs and attitudes -- positive evaluations of women conforming to traditional gender roles -- it appears men are considerably more likely to hold _hostile_ attitudes, or negative evaluations of women that violate gender roles (Barreto & Ellemers, 2005; Becker, 2010; Sibley et al., 2007). Hostile attitudes are the kinds of attitudes Fricker claims are operative in testimonial injustice. Men may reliably fail to gain information from women in a way that counts as an epistemic disadvantage if they systematically discount women who, in some sense, violated gender norms by speaking. But, the gendered differences in explicit prejudice could only explain a reliable pattern of epistemic disadvantage if women already have access to knowledge, evidence, or reasons that men don’t. In the cases Fricker uses to illustrate the phenomenon, the knowledge that men might have gained had they adequately trusted a woman’s testimony is knowledge they could nonetheless gain by some other means. If men are reliably
prejudiced, and reliably undervalue women’s credibility, we’ve certainly located an interesting claim about a gendered failure with epistemic consequences. There are some claims about disadvantage that might be developed from Fricker’s account -- e.g., men reliably fail to gain evidence from women’s testimony -- but these claims aren’t specific enough to support the idea that women have better judgments in any particular domain. Men aren’t cut off from information in virtue of its content, but in virtue of how it’s delivered. This approach is promising for the radical project only if we can also show that sexist discrediting leads to gendered differences in knowledge or understanding. Other attempts to account for men’s epistemic disadvantage might compliment this approach. One such approach appeals to epistemic vice.

3.4.2 - Vicious Men

Men may have an epistemic disadvantage if they are reliably willfully ignorant or epistemically vicious, and more so than women. José Medina also appears to make an argument along these lines when he claims that members of socially privileged groups tend to be actively ignorant. As previously states, Medina claims privileged agents may be reliably arrogant, (epistemically) lazy, and closed-minded (Medina, 2013, pp. 39–40). Moreover, he speculates these are due to the privileged agent’s motivation to not know (e.g., 2013, pp. 34, 36). Presumably, Medina is describing a refusal to accept the moral consequences of participating in oppressive social practices that benefit them (Baldwin, 2013; Moody-Adams, 1994; see also Ortega, 2006; Pohlhaus, 2012). The claims about epistemic vice and willful ignorance could be evaluated independently; however, it’s difficult to find any author claiming a privileged group is guilty of one flaw and not the other.
It may be the case that willful ignorance always involves some form of epistemic vice (Mason & Wilson, 2017) but, in theory, agents could have epistemic vices without willful ignorance. Willful ignorance involves an agent avoiding what they should and could do, and because of some motivation to maintain their self-image or their current way of life. Medina’s vices seem to involve some form of willful ignorance. The arrogant protects their self-image against threatening evidence, and the lazy person fails to do what they know they should because it’s easier not to try. The avoidance of ideas involved in closed-mindedness seems to typically involve arrogance, laziness, or the fear that one should change their mind. The etiology of gendered differences may not involve motives to avoid uncomfortable truths. For example, men and women may simply be socialized differently. But it’s difficult to argue the operation of these vices doesn’t involve some element of willful ignorance.

Regardless of whether epistemic vice and willful ignorance are distinct, there’s insufficient empirical support for an argument that men suffer from these problems more than women. Men may be more likely to overestimate their intelligence (Furnham, 2001). But arrogance may involve more than a simple overestimation of intelligence. It may involve patterns of interpersonal activity (Tiberius & Walker, 1998). A study of 2,653 Chinese students showed no gender differences in intellectual curiosity (Huang et al., 2010), and a study of 1,407 young adults suggests no gendered difference in open-mindedness (Shimai et al., 2006). Ultimately, Medina denies that we can reliably judge an agent is epistemically vicious from their social identity (e.g., Medina, 2013, p. 45), but defends the idea there is a close connection between these vices and social privilege. He suggests these vices are constitutive of oppressive social practices (e.g., 2013, p. 36). In this way, gendered differences in epistemic vice might be better understood as vices affecting specific domains of social judgment. It seems plausible that
privileged groups are motivated to maintain unjust social practices. If membership into social categories like gender are, in some way, constructed by oppressive social practices, then men might be expected to be epistemically vicious when considering matters of sexism.

In societies structured according to gender hierarchies, we might expect men to be more arrogant than women, and to be particularly arrogant towards women. Men might also regularly lack curiosity about social practices that benefit them, and to be particularly closed-minded when it comes to anti-sexist ideas that challenge their social privilege. But we shouldn’t expect women to not have these vices: if a woman enjoys social privilege in virtue of race, ethnicity, or economic status, she might also have predictable patterns of vice (Ortega, 2006). Though the vicious attitudes and practices of some socially privileged women might be similar at some level of description to those regularly found in men, these vices would hamper their epistemic success in different ways -- closing them off to different people and different ideas.

We can start to see promise here for developing a radical claim about women’s epistemic advantage. If oppressive social practices inform what it means to be a man, men may more regularly have attitudes that distort their abilities to attend to and process information that challenges their unjust social privilege. The claim is radical insofar as men’s judgments are said to be systematically compromised over a considerable epistemic domain. In some sense, a great deal of their social judgments may be predictably biased. These include judgments about whether some interpersonal interaction is sexist. But men may also routinely struggle to notice when some social arrangement affords them social privilege. Women may be considerably less likely to have the vices that hamper men’s judgments. Perhaps they lack the kind of masculine socialization that encourages arrogance, or they lack the social incentives to ignore men’s patterns of privilege. With the right explanation for vice (or lack thereof), the claim could also be
radical for implying an epistemic chasm. For example, men could have more epistemic success without the incentive of preserving their social privileges, but individual men can’t simply decide not to have these incentives. These privileges and incentives are granted by oppressive social practices that, as individuals, they have little control over. Moreover, if the very agents who could correct men’s epistemic vice are systematically discredited or silenced by these vices, or oppressive social practices generally (§3.4.1), this chasm would pose an even greater threat to moral or political progress. Ultimately, the amelioration of men’s epistemic disadvantage would require changes to society itself. Until the elimination of sexist oppression, we should expect even well-intentioned men to have distorted judgments about sexism.

If there is a close relationship between epistemic vice and social privilege, we may be able to recover radical claims about women’s epistemic advantage in the spirit of the projects abandoned by Standpoint theorists. But, it’s important to note, the domain of women’s epistemic advantage isn’t equally as large as the domain of men’s epistemic disadvantage. It would be inaccurate to move from an account of men’s epistemic disadvantage to the claim that women are, in some sense, experts on sexism or to the claim women enjoy special powers of judgment. While a woman may be far less likely to be burdened by vice than most men when making judgments about the social practices that harm them, the same woman may be burdened by vices that systematically distort their judgments about the general nature of sexism and the moral/political change needed to ameliorate oppression (see §3.3.2). These considerations may seem to undercut the radical spirit of early Standpoint claims; nonetheless, it would be sufficient epistemological justification for Standpoint’s methodological claims: that certain lines of inquiry ought to start from the lives of marginalized groups.
It would be premature to claim we’ve recovered Standpoint’s radical epistemological commitments by accounting for women’s epistemic advantage in terms of men’s vice. The success of this approach depends on establishing the claim that certain vices are regularly the product of social privilege and, further, establishing that men are more often than not vicious because of this regularity. These might be easy claims to establish in the context of a society where men are routinely willfully ignorant. In this context, men could and should have some awareness of sexist social practices, yet they avoid a full recognition of women’s oppression through vicious epistemic practices. They are motivated to retain their social privileges. This type of accounting for social privilege, vice, and willful ignorance is more commonly seen in explanations for why white Americans struggle to understand racist oppression (e.g., Baldwin, 2013; Eddo-Lodge, 2018). In some social circumstances, it may be true that men (or privileged groups generally) suffer from ignorance motivated by self-interest (Wieland, 2017, p. 7). This could explain why well-intentioned men still have unreliable judgments about sexism (e.g., Jones, 1999). But, if willful ignorance always involves morally vicious motivations (Wieland, 2017, p. 7), then this explanation would commit us to the view these so-called well-intentioned men are epistemically disadvantaged because they’re not really committed to anti-sexist values and practices.

Still, we can explain why well-intentioned men are at an epistemic disadvantage without denying they are sincere in their anti-sexist commitments. We can do this with an approach that unifies the best explanations for gendered epistemic difference.
3.4.3 - Vicious Societies

In the previous chapter, I argued social practices can render agents morally ignorant. Even an agent who is epistemically virtuous, and who is concerned about others’ moral interests, can falsely believe they are morally justified in participating in immoral or unjust social practices. As I’ve been using the term, “social privilege” denotes the material or psychological benefit one group gains from the social practices explaining the oppression of another group (Cudd, 2006, p. 25). These oppressive practices frequently include practices for conceptualizing and (mis)understanding social phenomena (Bonilla-Silva, 2006; e.g., C. Mills, 2007) -- enabling agents to believe their participation in oppressive practices is morally justified. Oppressive practices may also include practices for justifying dismissing testimony and discrediting testifiers (e.g., Dotson, 2011, 2014; Fricker, 2007) -- preventing agents from seriously considering challenges to their misconceptions and misunderstanding.

These practices may encourage socially privileged groups to develop dispositions that look like Medina’s vices. The difference is that social practices could lead agents to act in morally troubling ways while acting on rationally held beliefs. Even beliefs we can clearly identify as sexist prejudices may be rational to believe if the relevant beliefs are the best ones to have in light of the evidence (Begby, 2013; Gendler, 2011). A “well-intentioned” man might discount a woman’s testimony about subtle sexual harassment because he believes that sexual harassment has to look a particular way, and he believes women are prone to dramatically exaggerate events. When he gives her his reasons for believing she misinterpreted events, it would appear he’s arrogant -- especially if we already assume she has an epistemic advantage here. But, absent the recognition she has an advantage, he may have more reason to doubt her than to take her testimony at face value. Moreover, he may appear closed-minded when he
maintains his false beliefs after she tries to correct his misconceptions. Yet, if he maintains the beliefs that are best supported by his evidence after giving her point of view serious consideration, he may actually be practicing open-mindedness (cf. Kwong, 2015, 2017).

Sexist social practices, and the social privileges they afford to men, may explain gendered patterns of epistemic vice if many men are motivated to protect their social privileges (§3.4.2). But a more holistic appeal to social practices can better ground Medina’s observations that socially privileged agents are susceptible to willful ignorance while also explaining the appearance of epistemic vice in men who are genuinely trying to be anti-sexist. A man might appear to be engaged in motivated reasoning (Kunda, 1990) if his epistemic activity reliably leads him to self-serving false beliefs. Nonetheless, his epistemic activity need not be motivated by a vested interest in avoiding inconvenient moral facts (e.g., Wieland, 2017, p. 7). Sexist practices of conceptualizing social life, and sexist epistemic practices, can interfere with men’s abilities to rationally correct their moral ignorance even when they are attempting to do better. Moreover, sexist social practices can explain the prevalence of cognitive and affective biases that lead to epistemic injustice (§3.4.1). If women’s testimony is systematically undervalued, women would face considerable difficulty in correcting the sexist beliefs and practices that interfere with men’s ability to correct their ignorance.

An account that explains men’s epistemic disadvantage by a holistic appeal to sexist social practices is able to include for the insights of the bias- and virtue-based proposals while extending claims about men’s epistemic disadvantage. As long as some set of men’s epistemic successes are hampered by sexist social practices, and correcting men’s epistemic failings requires social change, we’re able to retain the radicalism of the claims recovered from the vice-based account. But, it might be objected, women may be just as likely to internalize sexist social
practices (e.g., Jost & Banaji, 1994). If this objection is right, then this holistic approach to explaining epistemic difference may fail to support the most basic claim there are differences between men and women. By following this line of inquiry, I argue the inability of this account to support the strongest claims about gendered epistemic difference might be better understood as a feature of the account and not a bug.

Very few Standpoint theorists were ever under the illusion that all women had an epistemic advantage merely because they are a woman. Many qualified their claims about women’s epistemic advantage with the recognition that many women internalize the beliefs and practices implicated in their own oppression. Nonetheless, while these beliefs and practices characteristically harm women, women are be more inclined to doubt the legitimacy of these practices (see §3.3). I argued we have reason to doubt that women’s experience of sexist oppression grants them more reliable judgments about sexism generally (§3.3.2). In spite of that, these experiences of harm can afford women distinctive reasons to doubt the system of beliefs that rationally support sexist prejudices and justify sexist social practices. In the previous chapter, I argued that agents rendered morally ignorant could not rationally correct any single false moral belief if it were rationally supported by a coherent system of bad beliefs (§2.2.3). For such a structurally ignorant agent, correcting particular false beliefs requires them to revise the entire system of beliefs that support them. Without strong reasons to do so, they have no epistemic obligation to perform such a demanding task. Women’s experiences of oppression may give them strong reasons to doubt the legitimacy of the systems of belief that justify sexist oppression, and encourage them to attempt to revise all they thought they knew about gender and social relations. If men characteristically benefit from sexist oppression, they would rarely have occasion to form strong doubts about sexist systems of belief. Thus, men would be
characteristically epistemically disadvantaged insofar as they are far less likely than women to be able to rationally correct their false, sexist beliefs.

Women might be said to have an epistemic advantage due to properties that commonly and characteristically hamper men’s ability to have reliable judgments about sexism. While bias and vice may play a role in explaining men’s epistemic disadvantage, these properties are explained by sexist social practices. A more holistic approach to explaining this disadvantage by appeal to social practices enables us to explain more cases of men’s recalcitrant ignorance on matters of sexism. While the same social practices do affect women’s epistemic success, the fact they are distinctively harmed by these practices affords them greater opportunity to correct for their pernicious influence on their judgments. The claim about women’s epistemic advantage defended here is radical in terms of the scope of the domain of judgments implicated, and in terms of implying an epistemic chasm between men and women that threatens moral and political progress. Moreover, it provides an epistemological justification for the methodological commitments that have since been the focus of Standpoint theorists. Plainly, inquiry on the nature of oppression should start from the experiences of oppressed people. This account gives a novel explanation of the epistemic value of those experiences, yet may raise new challenges for theorists developing methodological projects.

3.5 - Conclusion

In conclusion, I have argued it is possible to maintain a radical claim about women’s epistemic advantage: women may have more reliable judgments than men. This advantage may bring into question a large domain of men’s judgments, and men may be unable to gain more
reliable judgments on their own. I surveyed a number of approaches to grounding such a claim about gendered epistemic difference and argued they fail to explain any epistemic difference, fail to implicate a large domain of judgments, or fail to show men aren’t equally capable of sharing women’s epistemic successes. I reject claims that women have more reliable judgments by virtue of biological difference (§3.2) or differences in social experience (§3.3). I show how grounding women’s epistemic advantage in men’s epistemic disadvantage is promising, but reject simple appeals to men’s bias or epistemic vice (§3.4). I defend a radical claim about women’s epistemic advantage in virtue of oppressive social practices that explain men’s unreliable judgments about sexism and the distinctive barriers men face in correcting these errors.

The account of gendered epistemic difference I’ve developed here owes a great debt to the many insights made by Standpoint theorists, and resembles the more modest epistemological commitments developed in more recent methodological projects (e.g., S. Harding, 1992, 2015). Nonetheless, it is distinct from any radical account forwarded by Standpoint theorists, and constitutes a novel approach to developing Standpoint projects.
Conclusions: The Limits of Individual Moral Progress

Societies organized by socially unjust practices may stunt an individual’s capacity for moral improvement. These practices sustain, and are sustained by, unconscious psychological biases that we struggle to control (Chapter 1). Moreover, these practices may render us incapable of rationally correcting our prejudiced or problematic beliefs (Chapter 2). Together, biases and structural ignorance may explain why some political conflicts seem doomed. Depending on how social practices affect us, our ability to understand oppressive social practices may be hindered by our social identities (Chapter 3). I briefly review these arguments before considering avenues for future research.

Our biases and prejudices are not simply “bad habits” that we can correct through personal reflection, careful attention, and self-guided activity. Popular debiasing programs assume this self-help strategy is effective; however, I have argued we should doubt this strategy can help us ameliorate our implicit biases or the other biases that play a role in the reproduction and maintenance of socially unjust conditions. These bad biases are difficult to detect and monitor on our own. We may be able to prevent them being triggered in high-stakes decision-making contexts such as fielding job applicants, and we may be obligated to perform these interventions regardless of our ability to detect bias. Nonetheless, the empirical literature fails to give us practical and effective interventions to control the impact of bad biases in ordinary decision making and interpersonal interaction. Self-help strategies may be severely hindered by the persistence of sexist or racist representations in the environment and the social practices that sustain these representations. Direct interventions on these representations and practices may prove part of a more effective strategy for addressing the threat of our bad biases.
Our ability to identify that our social practices contribute to socially unjust conditions may be hindered by our upbringing. The beliefs and social practices we adopt early in life can bias our belief systems, and distort inquiry, in ways that keep us from rationally correcting our moral ignorance. I argued that agents rendered ignorant in this way are not responsible for failing to correct their ignorance -- against views that insist we’re always, in some sense, responsible for our moral ignorance. Not all moral knowledge is equally accessible to us, and we may appear to avoid or deny the truth without being motivated in blameworthy ways. Our beliefs may be rationally resistant to correction. Cases of ignorance like Megan Phelps-Roper’s may be rare. Nonetheless, beliefs and practices that maintain social inequality may saturate societies, and can mutually reinforce one another in ways that should cause us to doubt agents can easily find their way to the truth.

Last, I argued that ignorance explained by social practices might explain differences in epistemic success across social identities. Standpoint theorists were controversial for claiming that women have some epistemic advantage over men. These claims were radical for bringing into question large domains of knowledge (e.g., science and sociology), and for positing an unbridgeable epistemic difference between genders. I argue that attempts to ground women’s advantage in terms of some distinct power or capacity fail to support the radical claims about epistemic difference made by early Standpoint theorists. Nonetheless, I argue Standpoint’s radical claims can be recovered if women’s epistemic advantage can be explained in terms of men’s disadvantage. Sexist social practices might encourage men to be willfully ignorant, or to develop epistemic vice; however, men may still characteristically fail to understand sexism when they are making genuine attempts to understand. Sexist practices of conceptualizing social life, and sexist epistemic practices, may interfere with men’s ability to correct for structural
ignorance. If women have an epistemic advantage, it may be because their experience with these practices give them distinctive reasons to doubt the rationally coherent systems of belief that explain men’s ignorance.

When we consider the prospect of societal moral progress we may become pessimistic. Institutions, policies, biases, stereotypes, and false beliefs interact in complex ways to defend and reinforce the status quo. Though many have committed their lives to changing racist or sexist social practices, we may be deeply discouraged to witness prejudice and ignorance continuing to thrive in our communities. We may attempt to comfort ourselves with the idea that, if nothing else, we can do better as individuals; however, my arguments in these chapters challenge this notion. Cultural and social practices can constrain our ability to do, be, or know better.

In each chapter, I questioned optimistic views about our individual capacity for change: that we could work through our biases and ignorance with effort and the right attitude. My aim has not been to deny we are capable of bettering ourselves. It has been to account for factors that may limit our moral growth so that we may be more effective in our efforts to improve. If we fail to appreciate these limits we may be complacent with the illusion of personal growth and we may be too quick to vilify others for their failures. Yet, there are many questions left unanswered in my investigation. We may not be effective in our attempts to reduce our biases with the self-help strategy, but might there be moral value in using an ineffective strategy? If our biases are practically incorrigible, it is worth reconsidering what morality demands of us. Future research may find promise in investigating whether we can still live moral lives despite the ever-present moral hazards of automatic thought, feeling, and action.
Strategies for addressing bad bias through environmental changes may prove to be more effective than self-help strategies; however, greater empirical investigation is required to understand which environmental interventions work. We may discover reliable effective interventions that can be deployed in schools and businesses, and that have lasting effects when students and employers leave and are exposed to sexism and racism. But there may be significant social and political obstacles in implementing these interventions. Extant bias and diversity training may be ultimately pointless if an organization does not address the structural basis for discrimination and inequality: bias in their policies and social practices (Noon, 2018). This goes beyond the effects of social practices on an agent’s ability to apply the self-help strategy (see §1.4). These structural forms of racism and sexism can motivate agents to resist environmental interventions if employees interpret interventions as threats (e.g., §1.3.2). While hanging posters in the halls with counter-stereotype imagery may prove effective in a controlled experimental setting, the same intervention may prove counterproductive in especially homogenous workplaces. For example, employees at many engineering companies are predominantly white men. Even though companies may increase efforts at hiring women and black employees, they may struggle to retain them due to the prevalence of subtle sexist and racist attitudes in the workplace. Hanging posters of black and women engineers may only amplify these attitudes in the extant workforce -- becoming the butt of some shared joke. Ultimately, these complications highlight the fact that social inequity is not simply a matter of our bad biases. Social practices play a key role in the maintenance discrimination and disparity, and this may only be partially mediated by agents’ attitudes and ignorance.

If we fail to entertain the idea that social practices could distort our moral understanding, we may be more susceptible to their effects. But, how commonly are people victim to an
ignorance like Megan Phelps-Roper’s? When Michelle Moody-Adams rejected the inability thesis in the mid-1990s, American families were just starting to get personal computers and dial-up internet in their home. We might think that it could only be easier to correct our misconceptions now -- in an age where most Americans have instant access to the internet in their pocket. Discovering alternative points of view is only a Google search away, and we can engage in discussion with people who disagree with us instantly. Ironically, the internet may make it far more common for average people to fall victim to echo chambers that radicalize their belief systems.

Younger generations are increasingly getting their news and political information from online platforms that “personalize” the content delivered to the user. Machine learning algorithms track users clicks and view-times in order to deliver content to each user that is calculated to maximize clicks and view-times. These algorithms were originally developed to maximize advertising revenue for Facebook. If users were engaged for longer times, they could be targeted with more ads. But this marketing trick has some side effects. Users are already biased to engage with content that confirms their worldview -- statements they believe, from people they trust. Recommendation algorithms amplify these biases. Over the last decade, Facebook and YouTube have been under increasing scrutiny for shepherding young users to content produced by extremist groups -- groups with extreme beliefs and mistrust for non-believers. Before social media, agents belonging to echo chambers of political extremists, religious extremists, climate deniers, or anti-vaxxers may have become rationally trapped in these belief systems by particularly unfortunate circumstances. Now, these groups are seeing startling growth driven by recommendation algorithms. Because the operations of the algorithms are trade secrets, companies like Facebook and YouTube have been able to maintain they are not
responsible for this growth. Nonetheless, scholars are beginning to find ways to test publicly available information to demonstrate these trends (e.g., Ribeiro et al., 2020; Roth et al., 2020). Some thinkers may downplay the threat these algorithms pose by denying that people use social media networks to get their information, or by laying the blame on users for improperly exercising trust on an entertainment platform (cf. Rini, 2017). Yet, people passively acquire a great deal of information in their social media use even when they only intend to just catch up with friends: they cannot avoid being presented with trending stories or misleading political advertising. Moreover, Facebook’s strategies have been successfully adopted by Google and Amazon, making these three companies the largest advertising companies in history (Reyes, 2020; Vranica, 2020). Even people who shun social media platforms are likely to use Google to search for information. It’s public knowledge that agents can manipulate search results through “search engine optimisation” (SEO): burying stories that harm the reputation of their companies, or ensuring their brand is shown before others. However, initial studies have also demonstrated that Google’s algorithms can bias users' inquiry into political information (Grind et al., 2019; West, 2019).

My inspiration for studying structural ignorance came from social epistemologists studying social media networks (cf. Nguyen, 2018); however, I bracketed a serious consideration of social media in this dissertation. When I started my investigations, social media networks were not widely seen as a significant threat to social and political justice. Popular media outlets recognized that social media played a role in surprising political outcomes across the U.S. and Europe; however, they blamed individuals for misusing social media platforms. Since then, greater research has been dedicated to exposing the role algorithms play in shaping our social networks and the information we see. Moreover, investigation has exposed how political actors
have used these platforms to manipulate users for their own political gain (Berghel, 2018; e.g., Peters, 2018). In future research, I would like to investigate the role of social media (and other recommendation algorithms) in structural ignorance. If people today are more vulnerable to being rationally isolated by echo chambers than ever before, their inability to rationally correct their ignorance may result in intractable intergroup conflict. If we demonize affected individuals for failing to question their social practices we may be wrong in our assessments of their capabilities; moreover, our accusations against them may only backfire, and drive them deeper into their ignorance (see also §1.3.2).

In future research, I would also like to investigate the role of structural ignorance in other contexts of oppression. I’ve argued that our social identities may prove to have epistemological significance; however, my arguments about gendered epistemic differences. We may talk about oppression in a conceptually unified way: sexism and racism both involve an interconnected set of harmful and discriminatory social practices, but sexist and racist social practices operate in distinctive ways. While I suspect many insights developed here can be applied in the context of race, an account of “white ignorance” may be importantly distinct. Of note, racist practices commonly involve the physical segregation of people, which may be particularly conducive to the formation of echo chambers. Moreover, discriminatory legal policies and the generational effects of racial disparity may be interpreted to legitimize racist social practices, making white ignorance particularly resistant to correction. There have also been shifts in racist practices of conceptualization such racism can be more difficult to identify than sexism. An agent may believe they are engaged in an anti-racist practice that is actually racist. This is a distinguishing feature of color-blind racism (Bonilla-Silva, 2006), though might also be observed in performances of “white guilt” (Sullivan, 2014). Since these practices are justified by systems of
belief that include anti-racist attitudes, it may be particularly difficult for agents to have reasons to doubt their belief systems are biased.

While overly optimistic views about our capacity for moral growth may reinforce varieties of willful ignorance, overly pessimistic outlooks may fail to adequately account for the agency we have as individuals and in collectives. Ultimately, the psychological and social constraints reviewed here may encourage us to think more clearly about our ability to affect positive moral change in ourselves and in the world. In this imperfect world, personal moral development may require some degree of social change. As there are no simple solutions to systemic social problems, there may be no simple solution to leading a morally good life. Nonetheless, the stakes are too high to abandon hope. We may not yet know the best strategies for managing our biases and ignorance, but understanding our limits may be the first step to doing better.
**Bibliography**


Bateson, M., Callow, L., Holmes, J. R., Redmond Roche, M. L., & Nettle, D. (2013). Do images of “watching eyes” induce behaviour that is more pro-social or more normative? A field experiment on littering. *PloS One, 8*(12), e82055.


Batson, C. D., & Darley, J. (1973). From Jerusalem to Jericho: A study of situational and


Harvard University Press.


epistemic curiosity. *2010 IEEE 2nd Symposium on Web Society*, 258–263.


https://books.google.com/books?hl=en&lr=&id=8RxSY43LYj4C&oi=fnd&pg=PA13&dq=White+ignorance+Mills&ots=kWPCLOgeVF&sig=8t66ft1nx_c1yivGoKZud1hwXBs


Applications to addictive behaviors. *The American Psychologist, 47*(9), 1102–1114.


Reyes, M. S. (2020, December 3). Google, Facebook, and Amazon will account for nearly two-thirds of total US digital ad spending this year. *Business Insider.*

https://www.businessinsider.com/google-facebook-amazon-were-biggest-ad-revenue-winners-this-year-2020-12


https://philpapers.org/rec/SAUSAII


119–123.


Political Philosophy, 17(2), 202–227.


https://www.learner.org/series/discovering-psychology/the-power-of-the-situation/


overview for theoretical and philosophical psychology. Journal of Theoretical and


https://newrepublic.com/article/127179/time-france-abandon-laicite


Journal of Interpersonal Violence, 886260520975838.

Zablocki, B. D. (2001). Toward a demystified and disinterested scientific concept of
for objectivity in a controversial field (pp. 159–214). University of Toronto Press.


Zimbardo, Incorporated.